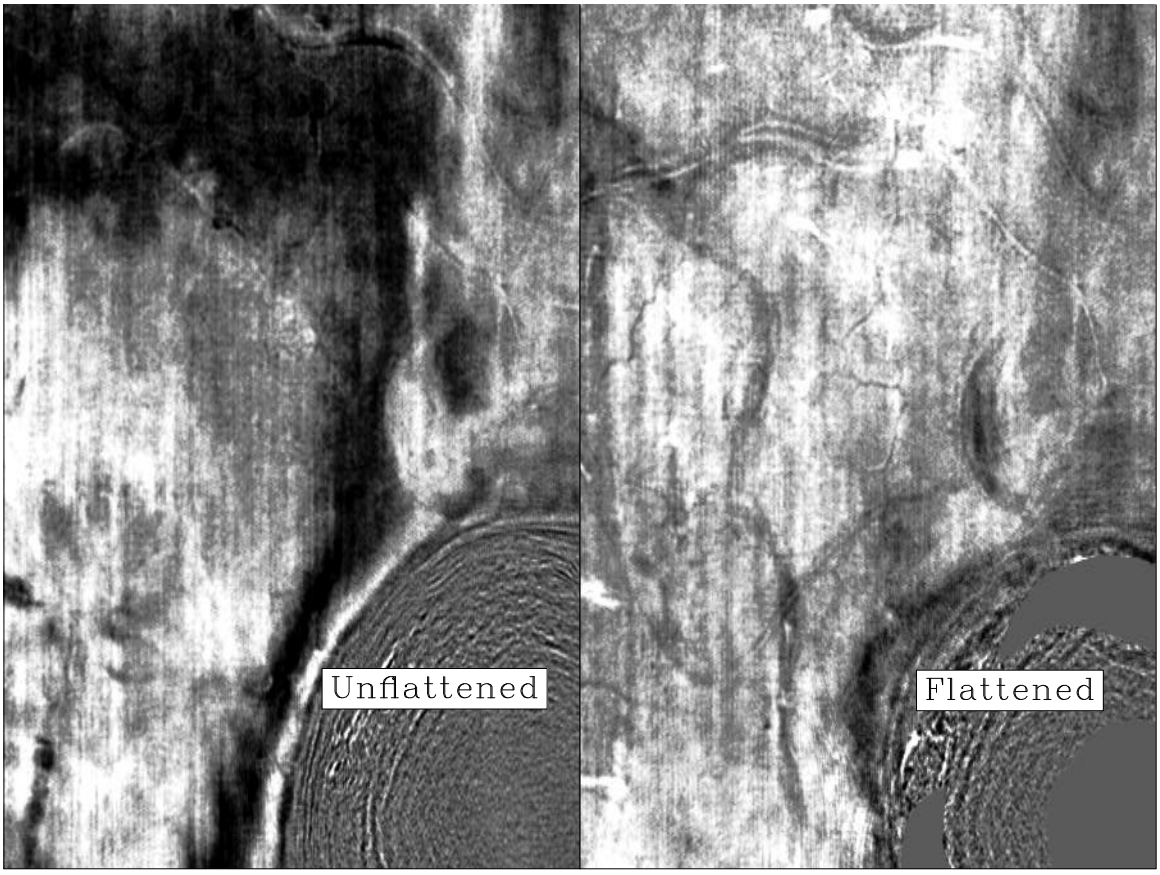
STANFORD EXPLORATION PROJECT

Gabriel Alvarez, Brad Artman, James Berryman, Biondo Biondi, Morgan Brown, Weitian Chen, Jon Claerbout, Marie Clapp, Robert Clapp, John Etgen, Christopher Liner, Jesse Lomask, Steven Pride, Daniel Rosales, Paul Sava, William Symes, Leon Thomsen, Alejandro Valenciano, and Ioan Vlad

Report Number 112, September 2002



Preface

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

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

CR denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. 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 figures. SEP discourages authors from flagging their figures as NR except for figures that are used solely for motivation, comparison, or illustration of the theory, such as: artist drawings, scannings, or figures taken from SEP reports not by the authors or from non-SEP publications.

Our testing is currently limited to LINUX 2.4 (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/research/redoc/>`.

[1] http://sepwww.stanford.edu/private/docs/sep112

[2] http://sepwww.stanford.edu/public/docs/sepdatalib/toc_html/

SEP–112 — TABLE OF CONTENTS

Velocity

Imaging

Optimization

Non-linear problems

Rock Physics

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 1–15

Ray based tomography using residual Stolt migration

Robert G. Clapp[1]

ABSTRACT

In complex areas, residual vertical movement is not an effective method to calculate traveltime errors for image domain tomography. By scanning over velocity ratios using residual Stolt migration, a different criteria for a coherent image can be defined, and a traveltime error approximated. The resulting traveltime errors are more accurate, therefore the tomography procedure is more robust than the more traditional methodology. Results are shown on a complex 2-D dataset.

INTRODUCTION

Depth migration is often necessary for complex structures, but requires an accurate interval velocity model. Estimating this velocity model is one of the essential problems in reflection seismology. One of the most common methods to estimate an interval velocity model is from ray-based reflection tomography after migration (Stork, 1992), which I will refer to as image domain tomography. In previous work (Clapp and Biondi, 2000; Clapp, 2001), I showed how to use angle gathers (Prucha et al., 1999; Sava and Fomel, 2000) in conjunction with downward continuation based migration (Gazdag and Sguazzero, 1984; Biondi and Palacharla, 1995) for back-projection. Post-migration reflection tomography is based on the fact that an offset gather (Kirchhoff) or angle gather (wavefield continuation) should be flat after migration. Deviation from flatness indicates a velocity error and can be converted into a travel time error and back-projected. Etgen (1990), among many others, pointed out that, in complex environments, looking at vertical moveout in gathers is not the optimal method to describe moveout errors. Biondi and Symes (2002) presented one alternative approach, constructing gathers where moveout is normal to an event.

Another approach is to use residual migration (Rocca and Salvador, 1982; Levin et al., 1983; Fomel, 1997) to find the best focusing velocity. Stolt (1996) and Sava (1999b,a) showed how to do residual migration for wave continuation methods. These methods allow scanning over slowness field ratios. Audebert et al. (1996) noticed that when scaling the slowness field by a constant, ray behavior is unchanged. Audebert et al. (1997) described a method of updating the velocity model by back projecting along the normal ray.

In this paper I take these works a step further. After performing downward continuation based migration, I use residual migration to find a smooth field of γ values that best focuses

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

the data. I then convert these γ values to approximate travel-time errors and back-project.

I begin by outlining a method of selecting my γ field. I then describe the approximations used to convert γ to $\Delta\mathbf{t}$. Finally, I show the procedure applied to a 2-D North Sea example.

BACKGROUND

Following the methodology of Clapp and Biondi (1999), I will begin by considering a regularized tomography problem. I will linearize around an initial slowness estimate and find a linear operator in the vertical traveltime domain $\mathbf{T}$ that relates change in slowness $\Delta\mathbf{s}$ with our change in traveltimes $\Delta\mathbf{t}$. We will write a set of fitting goals,

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\Delta\mathbf{s} \\ \mathbf{0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}, \end{aligned} \tag{1}$$

where $\mathbf{A}$ is our steering filter operator (Clapp et al., 1997) and ϵ is a Lagrange multiplier. However, these fitting goals don't accurately describe what we really want. Our steering filters are based on our desired slowness rather than change of slowness. With this fact in mind, we can rewrite our second fitting goal as:

$$\mathbf{0} \approx \epsilon\mathbf{A}(\mathbf{s_0} + \Delta\mathbf{s}) \tag{2}$$

$$-\epsilon\mathbf{A}\mathbf{s_0} \approx \epsilon\mathbf{A}\Delta\mathbf{s}. \tag{3}$$

Our second fitting goal can not be strictly defined as regularization but we can still do a preconditioning substitution (Fomel et al., 1997), giving us a new set of fitting goals:

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\mathbf{A}^{-1}\mathbf{p} \\ -\epsilon\mathbf{A}\mathbf{s_0} &\approx \epsilon\mathbf{I}\mathbf{p}. \end{aligned} \tag{4}$$

The trouble is how to estimate $\Delta\mathbf{t}$. Previously I have calculated semblance at various hyperbolic moveouts. I then picked the moveout corresponding to the maximum semblance. To calculate Δt I converted my picked moveout parameter back to a depth error Δz, which is converted into a travel time. Following the methodology of Stork (1992),

$$\Delta t = \cos(\alpha)\cos(\beta)v\Delta z, \tag{5}$$

where, α is the local dip, β is the opening angle at the reflection point, and v is the local velocity. This approach is effective in areas which are generally flat and have a sufficient offset coverage, but as shown in Biondi and Symes (2002) it runs into problems when these conditions aren't met.

A real data example of this problem can be seen in Figure 1. Note the "hockey stick" behavior seen at 'A'. If we follow the procedure of Stork (1992) we get unreasonably large travel time errors. Clapp (2002) shows that if we use these points for back projection, we get too large of velocity changes, which can lead to instability of the tomography problem.

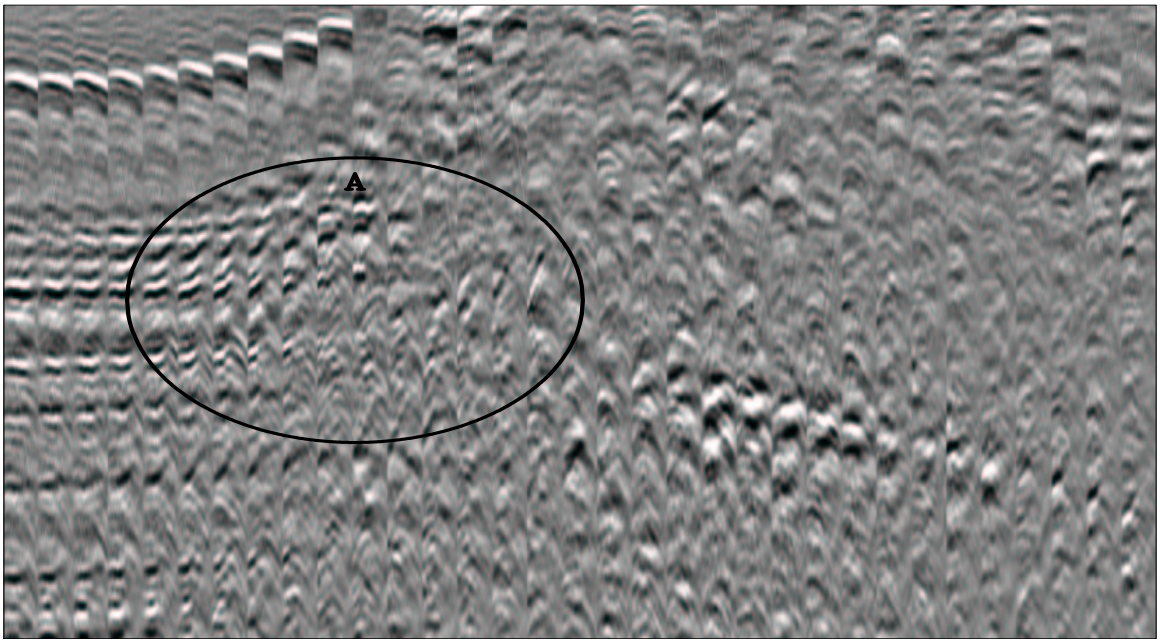


Figure 1: Every 5th gather to the left edge of a salt body. Note the coherent, "hockey stick" behavior within 'A'. bob1-gathers [CR]

METHODOLOGY

Creating a travel time error from residual migration involves several steps. My method is to first perform residual Stolt migration (Stolt, 1996; Sava, 1999b) on offset domain image gathers with a slight twist: minimize reflector movement among the various velocity ratios. I convert the different images to the angle domain (Sava and Fomel, 2000) and calculate the coherence of the zero moveout. I then set up a simple inversion problem to obtain a smooth estimate of γ for each model location. Finally, I convert this γ to an approximate traveltime error and back project it.

Residual Stolt migration

The main idea of Sava (1999b) is that we can do residual Stolt migration by defining the updated depth wave number k_z for various velocity ratios γ through the equation

$$k_z = \frac{1}{2}\sqrt{\frac{1}{\gamma^2}\frac{\left[4k_{z_0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|-\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]\left[4k_{z_0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|+\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]}{16k_{z_0}^2}-\left|\vec{k}_m-\vec{k}_h\right|^2} + \frac{1}{2}\sqrt{\frac{1}{\gamma^2}\frac{\left[4k_{z_0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|-\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]\left[4k_{z_0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|+\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]}{16k_{z_0}^2}-\left|\vec{k}_m+\vec{k}_h\right|^2}, \tag{6}$$

where

- k_{z0} and k_z are depth wavenumbers before and after residual migration,
- $\mathbf{v_0}$ the background velocity field,
- γ is defined as the $\frac{\mathbf{v_0}}{\mathbf{v_n}}$, and
- $\vec{k}_m$ and $\vec{k}_h$ are the offset and midpoint wavenumbers.

To make sure that residual Stolt migration behaves as we anticipate, I perform a test on the Marmousi synthetic. The top panel of Figure 2 shows every 15th common reflection point (CRP) gather of the Marmousi synthetic migrated with the correct velocity. Note how the CRP gathers are generally flat. The middle panel shows the same CRP gathers after migrating the data with a velocity 3% lower. The focusing has degraded and moveout in the gathers has increased. The bottom panel shows the result of performing residual migration with a γ value of .97. Note how, as expected, the image is better focused and the gathers are flatter.

Figure 3 shows an angle image gather for seven different values of γ ranging from .8 to 2.0. Note how we see the events move from curving down (left) to curving (up) as we increase γ. Unfortunately, the events also move as we change γ. We can eliminate this movement for flat events, and reduce it significantly for dipping events, by modifying our equation for k_z,

$$\begin{aligned} k_z = &\frac{1}{2}\sqrt{\frac{\left[4k_{z0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|-\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]\left[4k_{z0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|+\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]}{16k_{z0}^2}-\gamma^2\left|\vec{k}_m-\vec{k}_h\right|^2} \\ &+\frac{1}{2}\sqrt{\frac{\left[4k_{z0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|-\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]\left[4k_{z0}^2+\left(\left|\vec{k}_m+\vec{k}_h\right|+\left|\vec{k}_m-\vec{k}_h\right|\right)^2\right]}{16k_{z0}^2}-\gamma^2\left|\vec{k}_m+\vec{k}_h\right|^2}. \end{aligned} \tag{7}$$

Figure 4 shows the same image gather using equation (7). The reflector movement is generally eliminated from the gather.

Gamma selection

After performing the residual Stolt migration and converting to the angle domain, I am left with a volume of dimension $v(z,\alpha,x,\gamma)$, where α is aperture angle. From this volume we need to pick the best γ as a function of x and z. I also have the problem that even with the redefinition of the residual Stolt migration problem in Equation (7), events still have some movement at different γ's. For now I will ignore the movement problem on the theory that as long as we tend towards the correct solution, the best focusing γ will tend towards 1 and the amount of mispositioning at the best focusing γ will decrease.

For now I took a rather simple approach. I calculated the semblance for flat events at the different γ values. I then picked the best γ ratio at each location. I used this field as my data $\mathbf{d}$. I used the maximum semblance at each location as a weighting operator $\mathbf{W}$ to give more preference to strong events. I used a 2-D gradient operator for my regularization operator $\mathbf{A}$ and solved the inversion problem defined by the fitting goals,

$$\begin{aligned} \mathbf{0} &\approx \mathbf{W}(\mathbf{d}-\mathbf{m}) \\ \mathbf{0} &\approx \epsilon\mathbf{A}\mathbf{m}, \end{aligned} \tag{8}$$

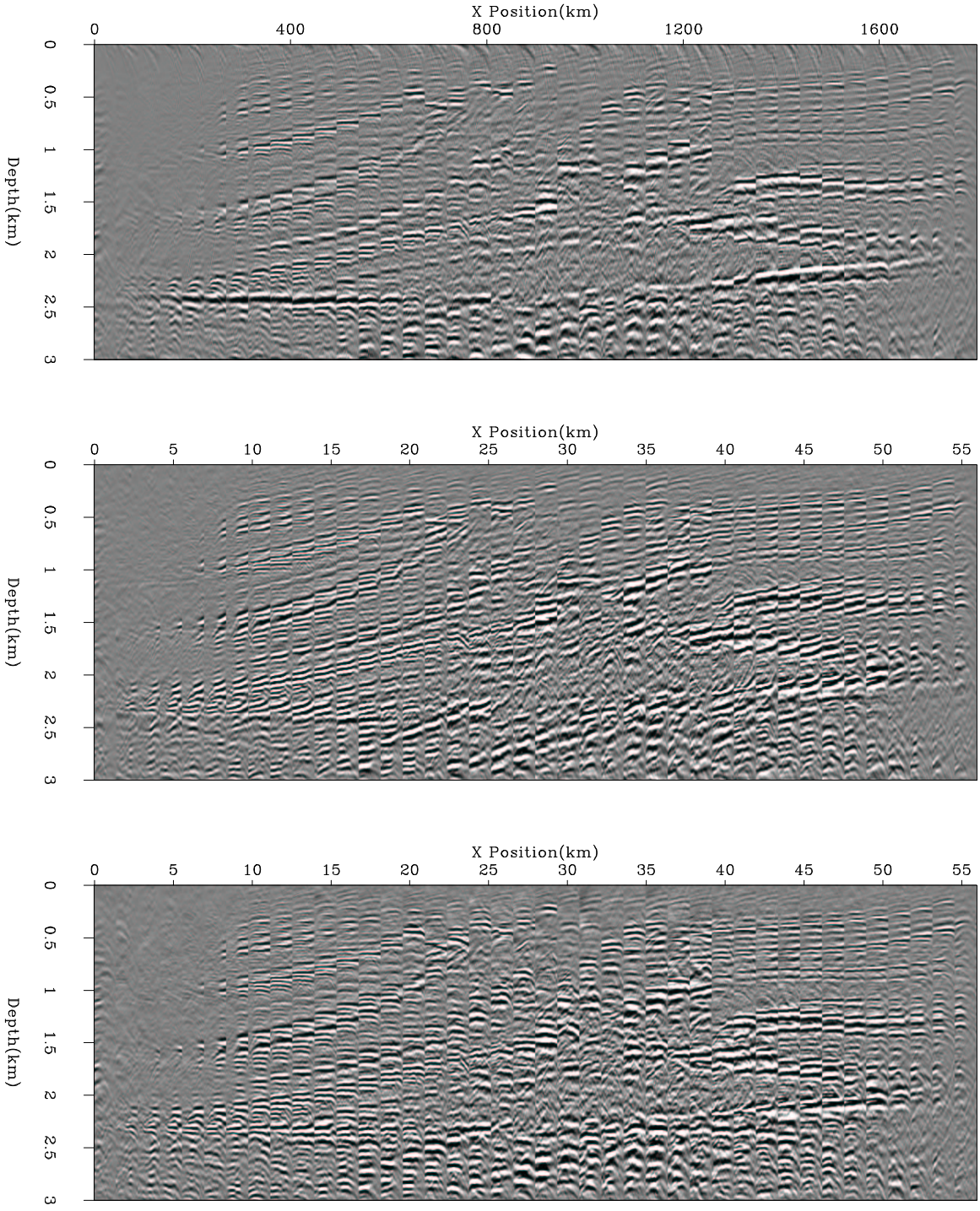


Figure 2: The top panel is the Marmousi synthetic migrated with the correct velocity. The center panel is the result of migrating with a velocity 3% less. The bottom panel is the result of residual Stolt migration with $\gamma = 97\%$. Note how the gathers are almost as flat as in the original migration. bob1-marm1 [CR,M]

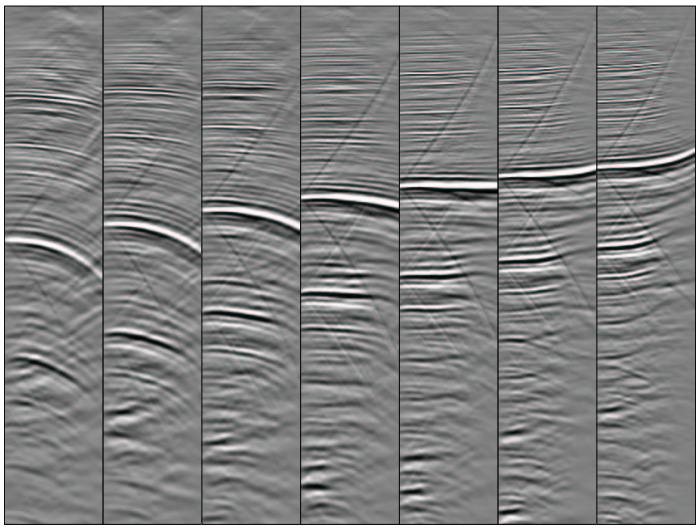


Figure 3: CRP gathers after performing residual Stolt migration with a γ value between .80 and 1.20. Note the significant reflector movement. bob1-moveme [CR,M]

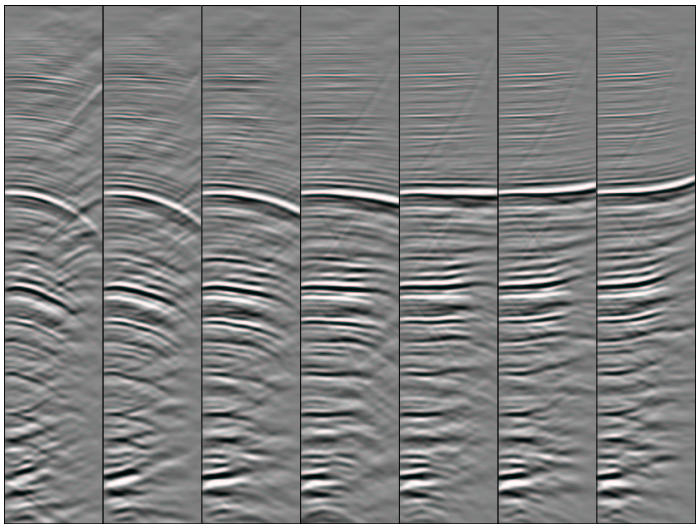


Figure 4: CRP gathers after performing residual Stolt migration using Equation 7. Note the lack of movement compared to Figure 3. bob1-nomove [CR,M]

where ϵ is the amount of relative smoothing and **m** is the resulting model. A better method, and a topic for future work, would be to calculate the semblance for a range of moveouts and do a non-linear search for a smooth γ function.

To test whether the method works I scanned over γ values from .95 to 1.05 on the migration result shown in the center panel of Figure 2. The left panel of Figure 5 shows the selected ratio, **m**, in fitting goals (8) and the right panel shows a histogram of the picked values. Note how we have generally picked the correct γ value ($\gamma = .97$).

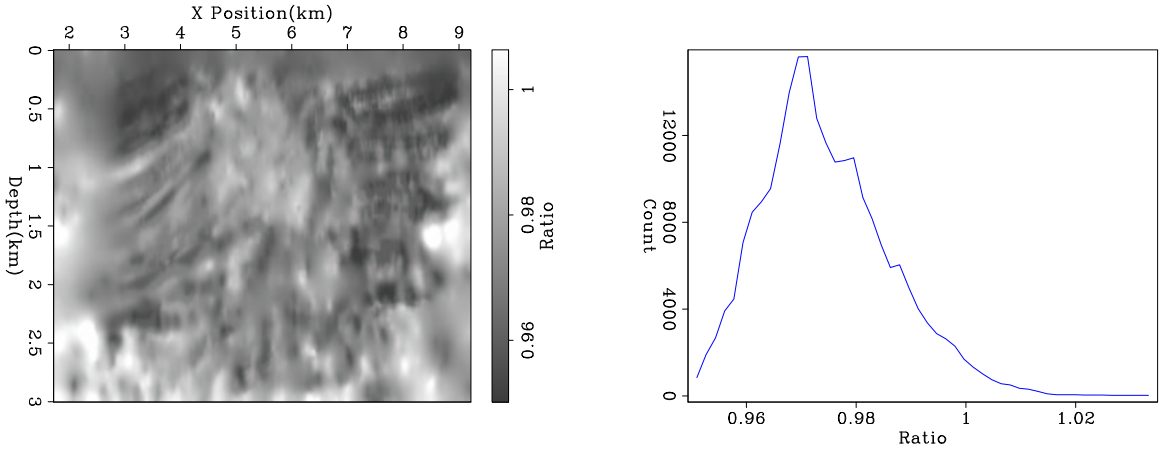


Figure 5: The left plot is the selected γ value using fitting goals (8). The right panel is a histogram of the picked values. Note the peak at approximately .97, the inverse of the velocity scaling. bob1-pick [CR,M]

Back projection

Converting the γ value into the $\Delta\mathbf{t}$ term needed for tomography is troublesome. Etgen (1990) showed that residual migration R_{mig} is the vector sum of residual normal moveout R_{nmo}, residual dip moveout R_{dmo} and residual zero offset migration R_{zoff} components. What we want to back project in image domain tomography is the R_{nmo} and R_{dmo} components. We can use simple trigonometry to convert a γ term to an approximate R_{nmo} term through

$$\Delta t(\mathbf{x},\alpha) = (\gamma - 1) * \left(\frac{1}{\cos(\alpha)} - 1 \right), \tag{9}$$

where $t(\mathbf{x},\alpha)$ is the traveltime to the surface of ray pair starting from **x** at the opening angle α. This approximation is not accurate, but has approximately the correct behavior.

EXAMPLE

To test the methodology I chose a 2-D line from a 3-D North Sea dataset. Figure 6 shows

- the initial velocity model (top-left panel),
- the initial migration (top-right), and
- a zoomed in portion of the model with three regions highlighted for future comparisons (bottom).

I performed one non-linear iteration of tomography using three different approaches. The top panel of Figure 7 shows the result of using vertical moveout as the basis for determining the time errors. The center panel is again the result of using vertical moveout for the time error calculation with additional constraints on what points are used for back projection and limiting the effect of bad moveouts (Clapp, 2002). The bottom panel is the result of using residual migration.

The area signified with 'A' shows the problems with using all of the data (top panel). Note how the gathers have tremendous, inconsistent curvature. When we discount this information (center panel) we get more reasonable gathers. Using residual migration to estimate $\Delta\mathbf{t}$ we can get the same, or better, gathers without throwing away a portion of the data. The problem with throwing away a portion of the data can been at 'C' and especially 'B'. We threw away the information that would help us flatten the reflector in order to avoid the problems seen in the top panel. Smply using vertical moveout analysis, it takes several non-linear iterations to achieve the same level of flatness at 'A' and 'B' that is seen with the first non-linear iteration using a residual migration measure.

Figure 8 shows every 15th CRP gather after five non-linear iterations. Note how the moveout in reflector at 'A' is virtually flat. Within valley structure, 'B', there is little remaining residual moveout. The most interesting location is 'C' where we are begining to see coherent events under the salt. Figure 9 shows the resulting velocity and image after five non-linear iterations. The salt top reflection is now clean. Note how the valley structure at 'A' is well imaged. At 'B' we can follow reflectors all the way to what appears to be the salt edge. On the top-left portion of the salt, 'C', we have gone from a jumbled mess (Figure 6) to being able to clearly follow reflectors. At 'D' we see a consitant, strong amplitude, salt bottom reflection. Finally, at 'E' we are begining to see strong events under the salt. Further improvement requires going to 3-D.

Another way to evaluate image improvement is to look at the γ values after successive iterations. Figure 10 show the γ values after zero to four iterations (left to right, top to bottom). The figure demonstrates that as we progress in iteration the γ value tends towards 1, indicating that the problem is converging.

CONCLUSIONS

Vertical moveout is inadequate method to characterize migration errors caused by velocity. By using residual Stolt migration more reasonable errors can be estimated and back projected. Early results indicate that the inverted result is more promising than simply using vertical positioning errors.

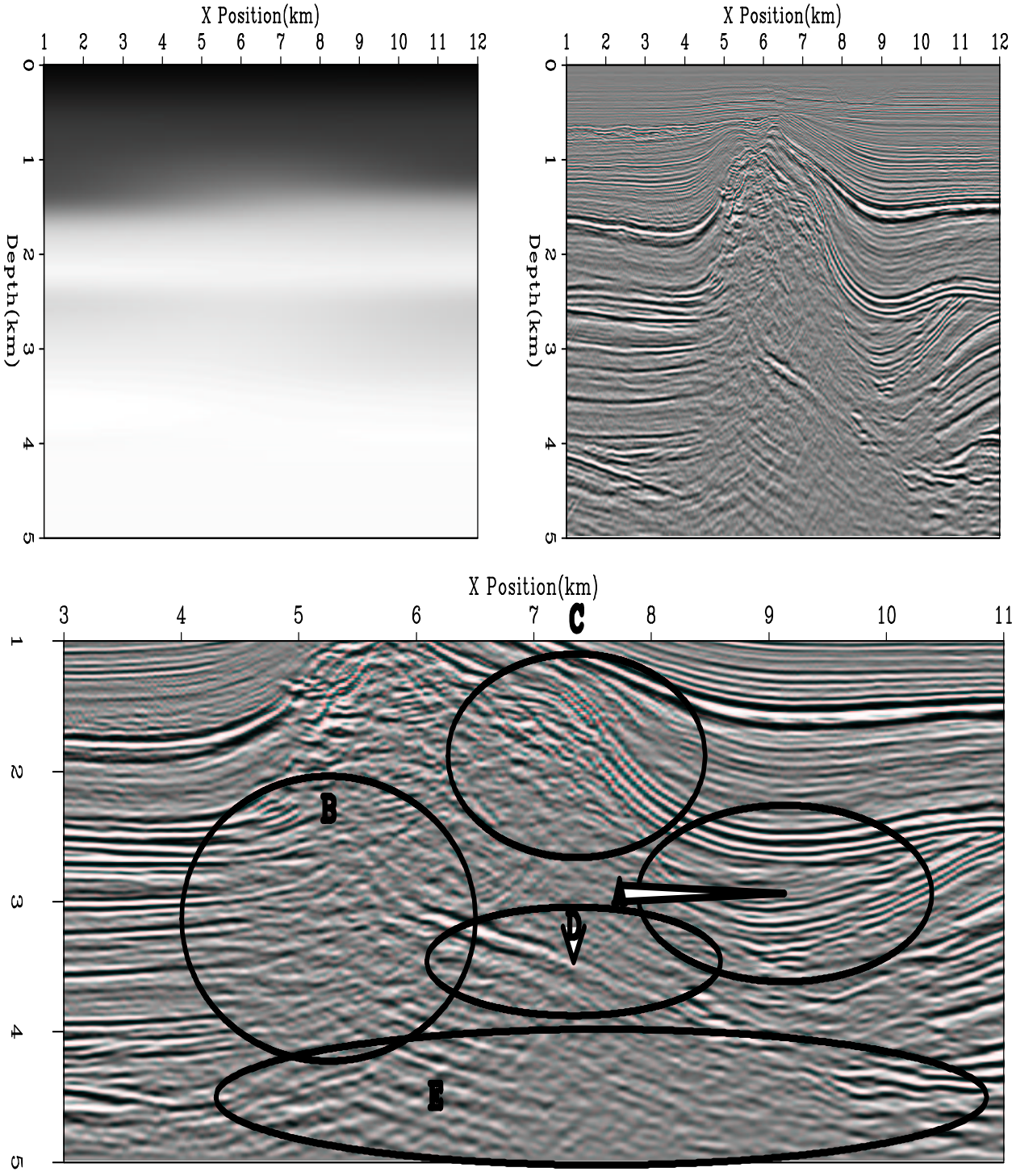


Figure 6: The initial model and migration. The top-left panel shows the velocity model and the top-right panel shows the migrated image using this velocity. The bottom panel shows a blow up around the salt body. 'A' - 'E' will be used later in the text for comparison. bob1-combo.vel0 [CR,M]

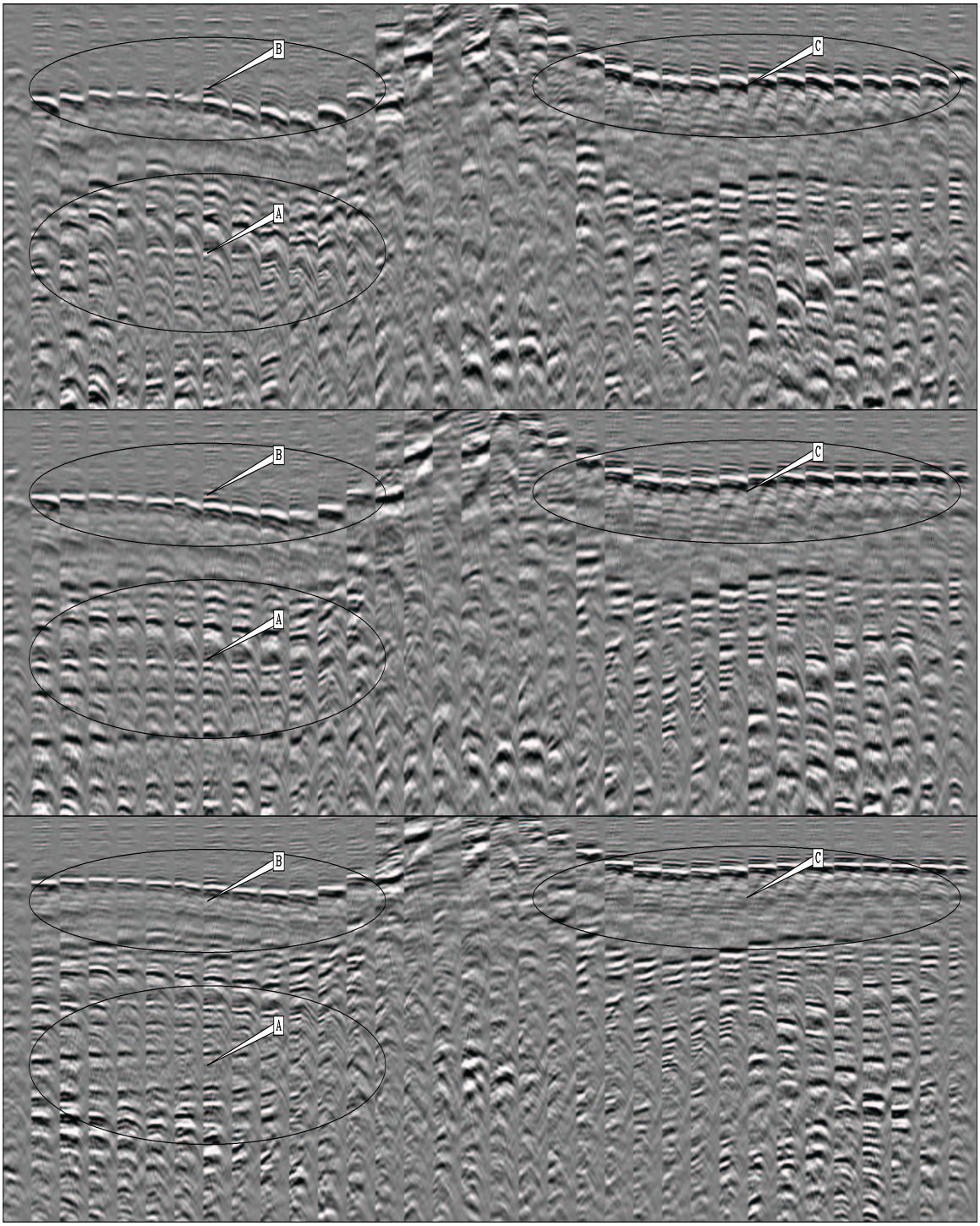


Figure 7: Each panel represents every 15th CRP gather between 1 and 4 km after one nonlinear iteration of tomography. The top panel is the result of performing tomography calculating $\Delta\mathbf{t}$ using vertical moveout. The center panel is using vertical moveout discounting data with significant moveout. The bottom panel shows the result of using residual migration as the basis of the $\Delta\mathbf{t}$ calculation. Note the improved flatness of the CRP gathers from top to bottom. bob1-gathers.iter1 [CR,M]

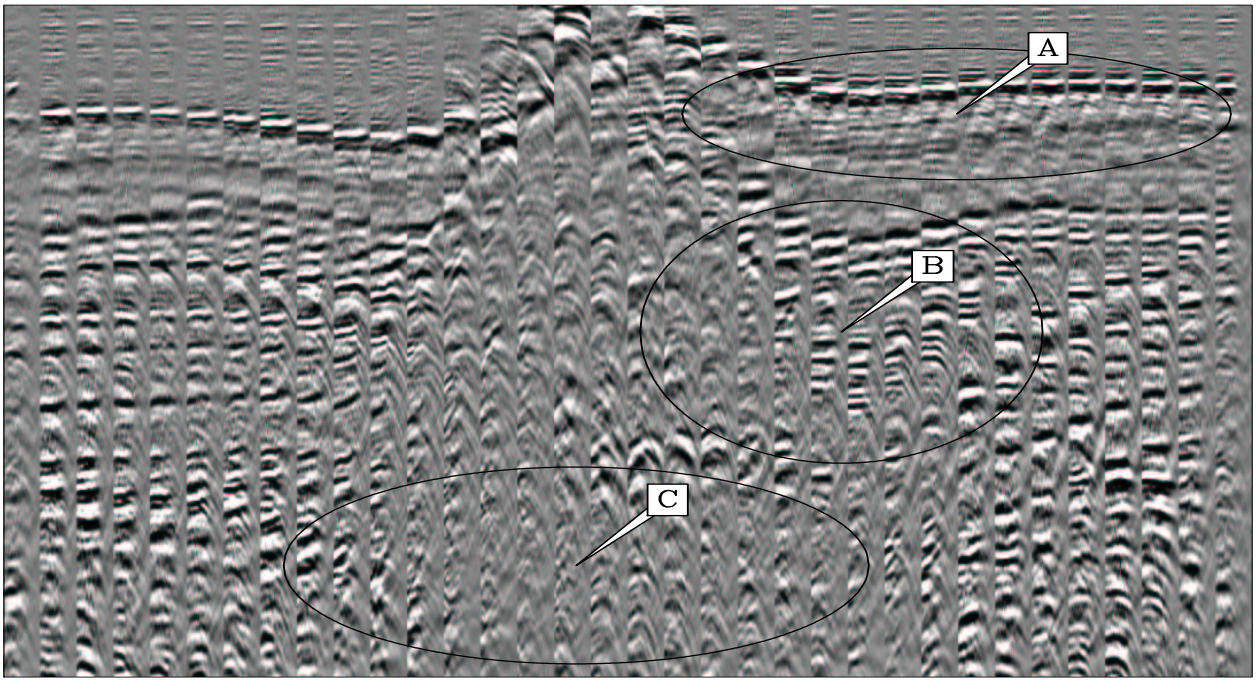


Figure 8: Every 15th CRP gather after five non-linear iterations. Note the flat gathers at 'A' and 'B' and the forming of coherent moveout below the salt at 'C'. bob1-gathers.final [CR]

REFERENCES

Audebert, F., Diet, J. P., and Zhang, X., 1996, CRP-scans from 3-D pre-stack depth migration: A powerful combination of CRP-gathers and velocity scans:, *in* 66th Ann. Internat. Mtg Soc. of Expl. Geophys., 515–518.

Audebert, F., Diet, J. P., Guillaume, P., Jones, I. F., and Zhang, X., 1997, CRP-scans: 3-D preSDM velocity analysis via zero-offset tomographic inversion:, *in* 67th Ann. Internat. Mtg Soc. of Expl. Geophys., 1805–1808.

Biondi, B. L., and Palacharla, G., 1995, 3-D prestack depth migration of common-azimuth data:, *in* 65th Ann. Internat. Mtg Soc. of Expl. Geophys., 1197–1200.

Biondi, B., and Symes, W. W., 2002, Transformation to dip-dependent common image gathers: SEP–**112**, 65–82.

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

Clapp, R. G., and Biondi, B. L., 2000, Tau tomography with steering filters: 2-D field data example: SEP–**103**, 1–19.

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

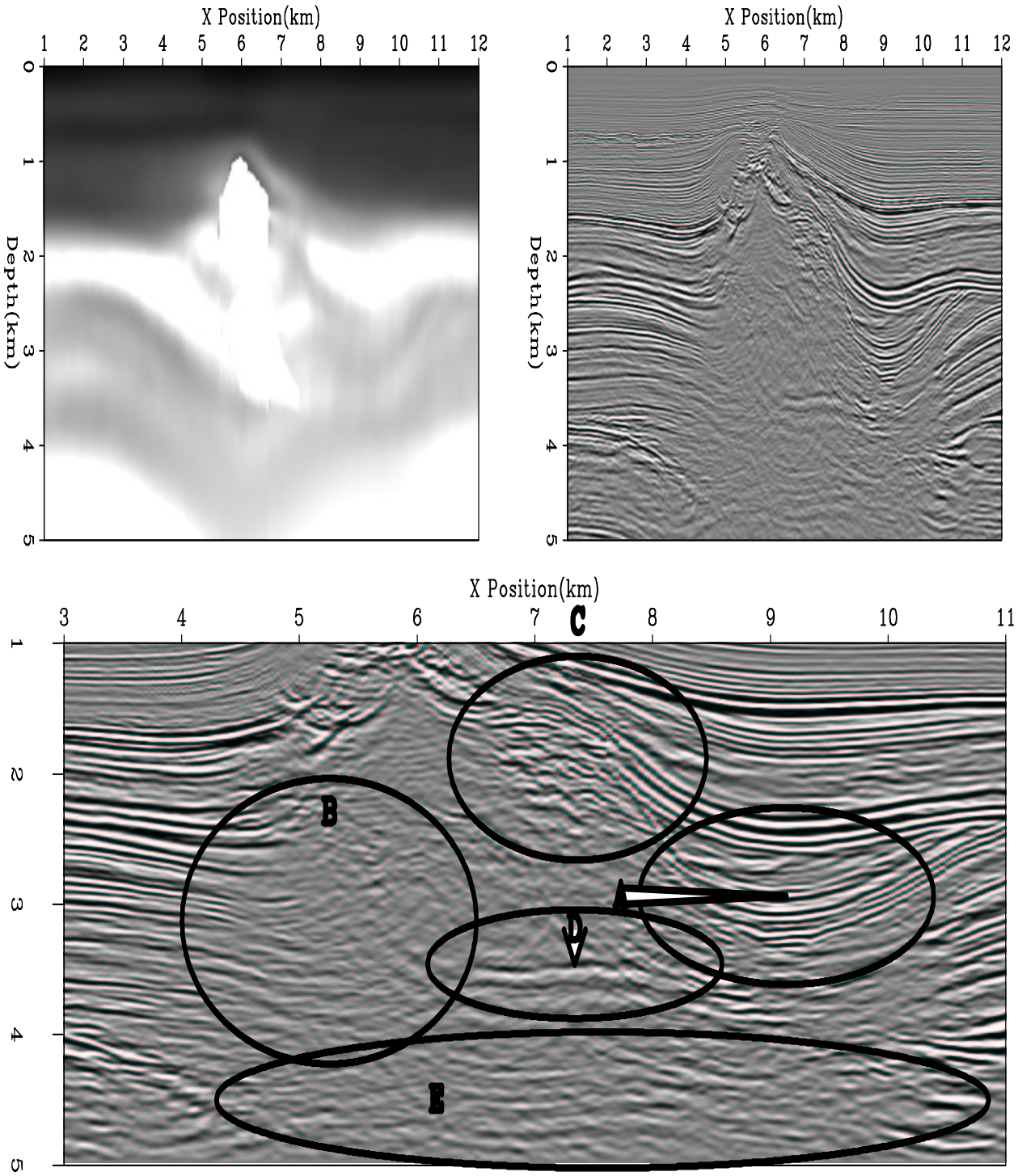


Figure 9: Data after five non-linear iterations of tomography with residual migration based moveout analysis. The top-left panel shows the velocity model and the top-right panel shows the migrated image using this velocity. The bottom panel shows a blow up around the salt body. Note how the valley structure at 'A' is well imaged. At 'B' we can follow reflectors all the way to what appears to be the salt edge. On the top-left portion of the salt, 'C', we have gone from a jumbled mesh (Figure 6) to being able to clearly follow reflectors. At 'D' we see a consistent, strong amplitude, salt bottom reflection. Finally, at 'E' we are begining to see the forming of fairly strong events under the salt. bob1-combo.final [CR,M]

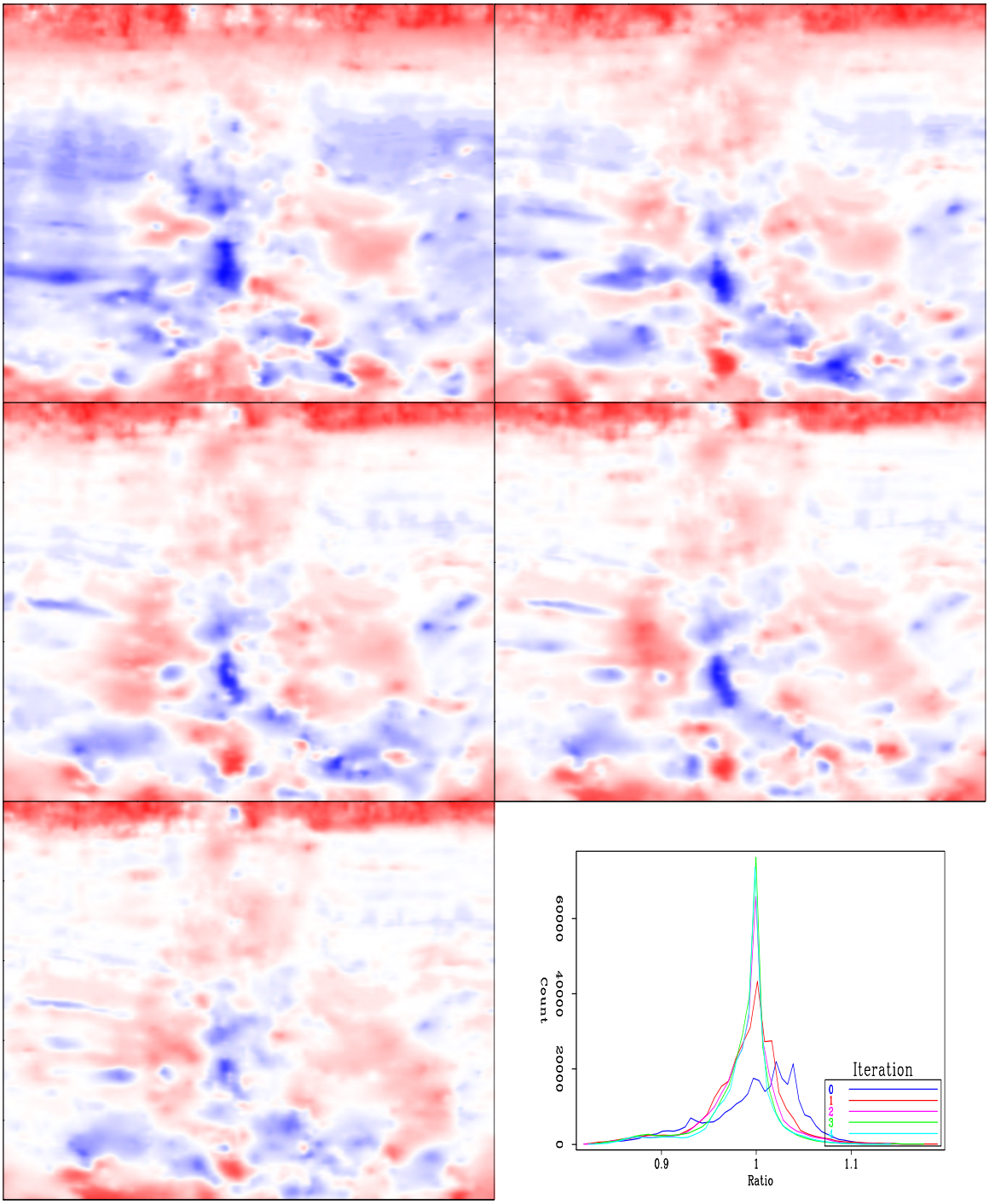


Figure 10: The residual moveout measure as a function of non-linear iteration. The top-left is the initial moveout, the top-right after one iteration, etc. The bottom-right panel shows a histogram for each of the γ maps. Note how the points cluster towards 1 as we progress in iteration. Note how the ratio tends towards 1 (no-moveout, white) as we progress. The red stripe on top is due to the severe early mute. The offset to angle transform has an edge effect that causes the data to curve up. As a result the best flatness is obtained with a very low γ value. These values are ignored in the inversion. bob1-ratios [CR,M]

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

Clapp, R. G., 2002, Dealing with errors in automatic velocity analysis: SEP–**112**, 37–46.

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

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

Fomel, S., 1997, Velocity continuation and the anatomy of prestack residual migration:, *in* 67th Ann. Internat. Mtg Soc. of Expl. Geophys., 1762–1765.

Gazdag, J., and Sguazzero, P., 1984, Migration of seismic data by phase-shift plus interpolation: Migration of seismic data by phase-shift plus interpolation:, Soc. of Expl. Geophys., Geophysics, 124–131.

Levin, S. A., Rothman, D., and Rocca, F., 1983, Residual migration: Applications and limitations:, *in* 53rd Ann. Internat. Mtg Soc. of Expl. Geophys., Session:S10.7.

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

Rocca, F., and Salvador, L., 1982, Residual migration:, *in* 52nd Ann. Internat. Mtg Soc. of Expl. Geophys., Session:S1.4.

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

Sava, P., 1999a, Short note–on Stolt common-azimuth residual migration: SEP–**102**, 61–66.

Sava, P., 1999b, Short note–on Stolt prestack residual migration: SEP–**100**, 151–158.

Stolt, R. H., 1996, Short note - a prestack residual time migration operator: Short note - a prestack residual time migration operator:, Soc. of Expl. Geophys., Geophysics, 605–607.

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

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 15–21

Short Note

Wave-equation MVA applied to 4-D seismic monitoring

Paul Sava, John Etgen, and Leon Thomsen[1]

INTRODUCTION

4-D seismic processing is gradually maturing as a technique able to aid time lapse monitoring of seismic reservoirs (Lumley, 1995; Biondi et al., 1996). However, many limitations hamper the ability of 4-D seismic monitoring to produce reliable results in complicated reservoir situations. One such example is that of multi-layer reservoirs where changes at deeper levels are masked by those that occur at the top reservoir. In these cases, only the top-most reservoir is analyzed and changes at the deeper levels are disregarded or at least treated as suspect.

Of particular interest is the case of reservoirs where the pressure configuration is such that gas is at the limit of release in solution (Kristiansen et al., 2000). Any drop in pressure, likely to occur during production, leads to gas release which results in substantial change in velocity. In these cases, the 4-D effects are mainly driven by the changes in acoustic velocity. For these reservoirs, 4-D seismic monitoring can be seen as a velocity analysis problem.

Biondi and Sava (1999) introduce a method of migration velocity analysis based on wave-equation techniques (WEMVA) which uses the changes visible in the entire seismic image to infer velocity information. Such a technique is ideally suited to deal with velocity-related 4-D changes observed over entire images, including the case of multi-layer reservoirs.

Traveltime-based MVA methods cannot be easily used to solve this problem for several reasons: the traveltime changes that occur over time are too small to be picked with enough accuracy; amplitude information, although very important, cannot be used and is, therefore, ignored.

WEMVA applied to 4-D problems has limitations as well. First, WEMVA can only handle the image changes due to perturbations of the acoustic velocity since our current implementation is based on the acoustic wave-equation. Second, WEMVA can only handle small velocity anomalies, due to the inherent Born approximation. This, however, is unlikely to be a problem for 4-D analysis since the image changes are smaller than a fraction of the seismic wavelet.

[1]**email:** paul@sep.stanford.edu,john.etgen@bp.com,thomsela@bp.com

METHODOLOGY

In its original formulation (Biondi and Sava, 1999), wave-equation migration velocity analysis relates a perturbation of the slowness model ($\Delta\mathbf{S}$) to its corresponding perturbation of the seismic image ($\Delta\mathbf{R}$). Mathematically, this relation can be expressed as the linear fitting goal

$$\mathbf{L}\Delta\mathbf{S} \approx \Delta\mathbf{R}. \tag{1}$$

$\mathbf{L}$ is the WEMVA operator that is constructed as a linearization of downward continuation operators involving the Born approximation (Sava and Fomel, 2002). We obtain the slowness perturbation $\Delta\mathbf{S}$ from Equation (1) by applying either the adjoint or the least-squares inverse of $\mathbf{L}$ to the image perturbation $\Delta\mathbf{R}$.

The critical quantity in Equation (1) is the perturbation of the seismic image $\Delta\mathbf{R}$. For the purpose of this equation, this is the known quantity and various techniques can be used to derive it.

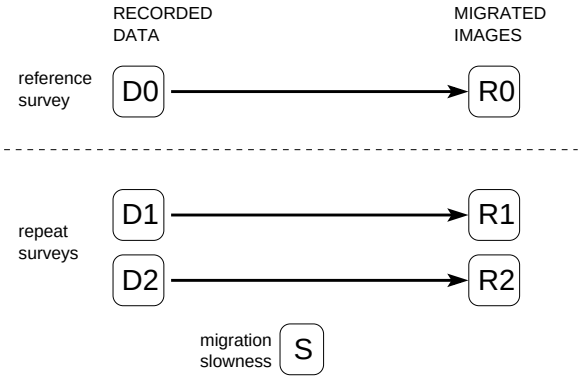


Figure 1: Different 4-D datasets imaged using the same slowness model produce different seismic images, from which we can extract image differences for WEMVA. paul1-4Dscheme [NR]

In 4-D seismic monitoring, the image perturbation is defined as the difference between the images at various acquisition times with respect to the reference image. For example, suppose that at time $t = 0$ we record a reference dataset $\mathbf{D}_0$ which is imaged with the migration slowness $\mathbf{S}$ to produce the reference image $\mathbf{R}_0$. At later times, repeat surveys produce new datasets $\mathbf{D}_1, \mathbf{D}_2 \ldots$ which are different from $\mathbf{D}_0$ and, therefore, reflect the changes in the reservoirs.

After imaging using the same slowness model $\mathbf{S}$, we obtain the images $\mathbf{R}_1, \mathbf{R}_2 \ldots$ which are different from the reference image $\mathbf{R}_0$ (Figure 1). The image differences or perturbations are obtained by simply subtracting the reference image from each of the repeat images. Once we have created the image perturbations $\Delta\mathbf{R}_1, \Delta\mathbf{R}_2 \ldots$, we can invert for slowness perturbation $\Delta\mathbf{S}$ using Equation (1).

EXAMPLE

We illustrate this technique with a synthetic model that resembles a typical producing reservoir in the North Sea. The model is depicted in Figure 2: the reflectivity on the left, and the reference slowness on the right. The model consists of several fractured horizontal reservoirs

which are in production. The reference slowness is smooth and not conformant with the stratigraphy. We assume that the reference slowness **S** is derived from the reference survey and that it perfectly focuses the reference data $\mathbf{D}_0$ to create the reference image $\mathbf{R}_0$.

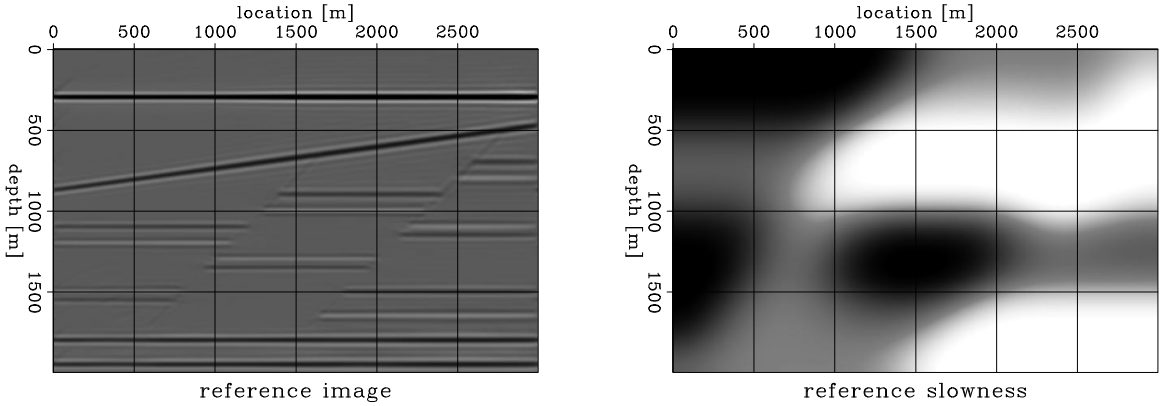


Figure 2: Synthetic model: reflectivity (left) and slowness (right). paul1-model [CR]

Figure 4 shows the slowness perturbations we introduce in the slowness model. For each of the two scenarios, we generate data using the same reflectivity (Figure 2) but different slowness models generated by adding the respective slowness perturbations to the reference slowness. We then image using the reference slowness to create repeat survey images. Finally, we subtract the reference image from each of these two images and obtain the image perturbations (a.k.a. the 4-D seismic data) depicted in Figure 3.

Figure 3 enables us to make two observations:

- Although the changes in the reservoirs occur at only two levels, the changes in the images occur at all levels underneath. This situation is common for 4-D seismic surveys. Typically only the top reservoir can be properly analyzed since the 4-D effects created by the deeper reservoirs are either masked or seriously shadowed by the top reservoir.
- Completely different changes in the reservoirs yield fairly similar perturbations of the images. Even for such a simple model, as the one we use in this analysis, it is really hard to visually analyze the image perturbation and distinguish among the two cases (Figure 3). In practice, this distinction is virtually impossible, and the only place where we can extract reliable information is at the top-most producing reservoir.

We address the ambiguity of the 4-D interpretation using WEMVA. Figure 5 shows the slowness perturbations obtained by the adjoint of the WEMVA operator **L** in Equation (1) applied to the image perturbation $\Delta\mathbf{R}$ in Figure 3. The two cases can be better distinguished now, although the information is not yet localized at the producing reservoirs.

The least-squares inversion result, shown in Figure 6, is much better focused at the reservoirs. Despite the inherent vertical smearing mainly caused by the limited data aperture, we can precisely indicate the location of the producing reservoirs, the sign of the slowness change, and even the relative magnitude of the change from one reservoir to the other.

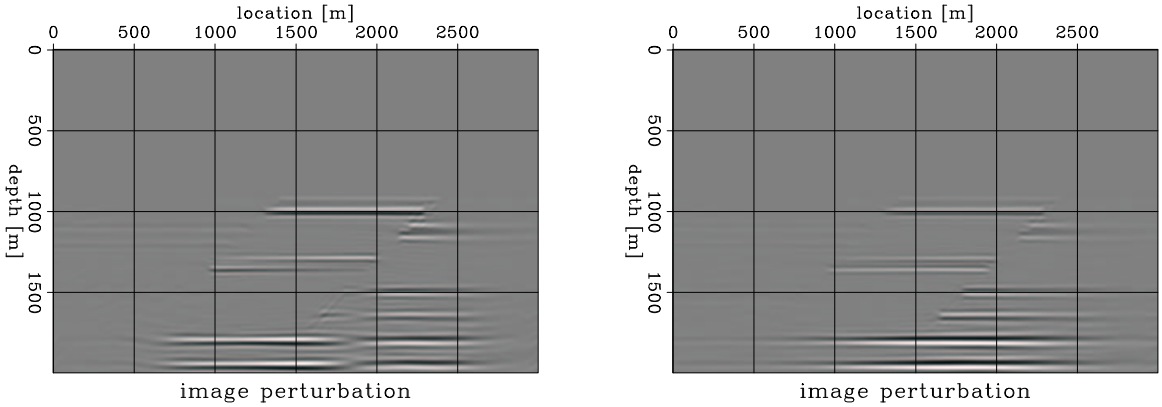


Figure 3: Image perturbation: scenario 1 on the left and scenario 2 on the right. paul1-dimag [CR]

DISCUSSION

The processing method outlined in this paper is mainly applicable to the situations when the 4-D effects translate into significant slowness variations, for example in cases where pressure changes lead to release of gas in solution and consequently to a drop in velocity. Furthermore, the method is strongly dependent on the quality of the recorded data and also on the quality of the 4-D pre-processing.

We must insure that our definition of the image perturbation is mainly a product of the slowness model perturbation. Much care needs to be taken to eliminate all acquisition differences between the repeat surveys and all processing differences of the different datasets. An ideal case consists of fixed acquisition (permanent water-bottom receivers, for example) and identical seismic processing.

Correct handling of amplitude data in migration is as important as in any method addressing reservoir-related properties. However, in this method we are mainly concerned with the differences between repeat images and not as much interested in their absolute magnitude. Therefore, this method is likely to be robust with respect to the accuracy of the more or less accurate migration amplitudes. This particular subject, however, requires careful further analysis.

CONCLUSIONS

We present an application of the WEMVA methodology to 4-D seismic data. If certain physical conditions are met, this velocity analysis method is capable of identifying the producing reservoirs, even for the cases of production from multiple levels.

4-D pre-processing remains an important component of the method. We need to insure that all image perturbations are not related to differences in acquisition and/or processing, but

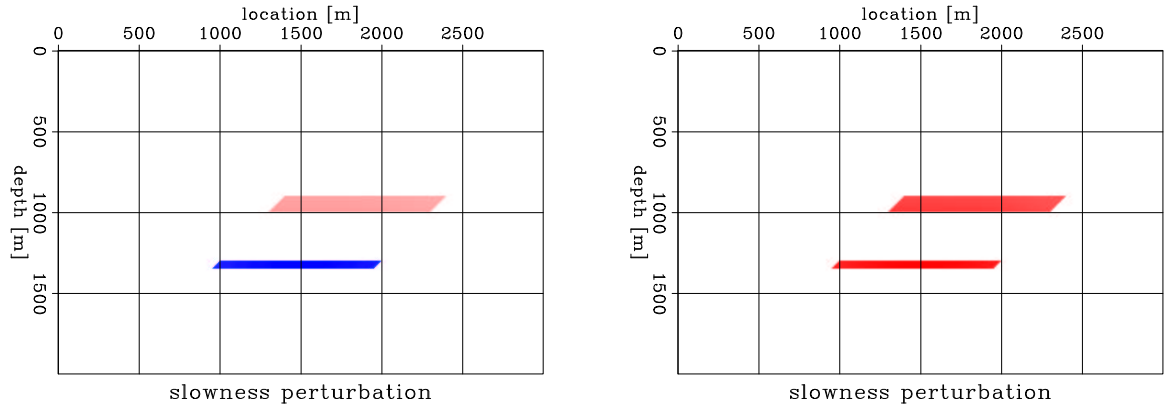


Figure 4: Slowness perturbation: scenario 1 on the left and scenario 2 on the right. paul1-dslow [CR]

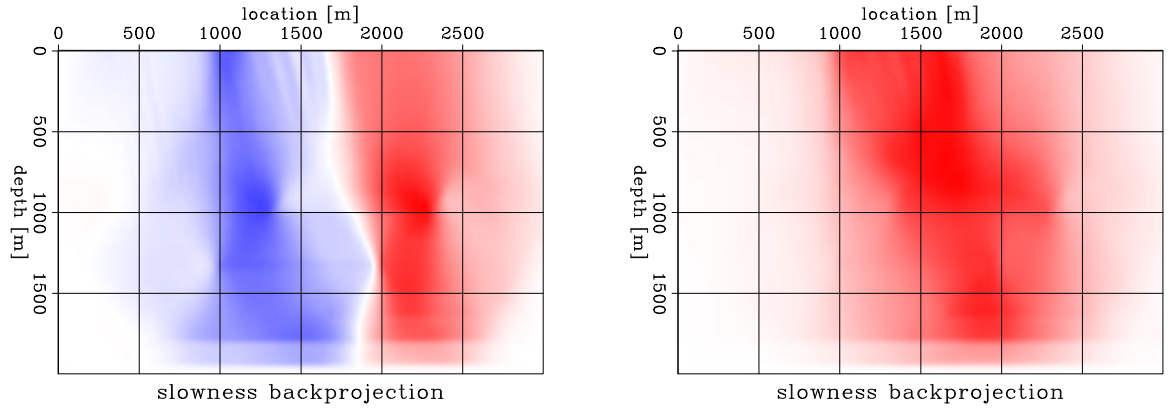


Figure 5: Slowness perturbation obtained using the adjoint of operator **L** in Equation (1). paul1-bslow [CR]

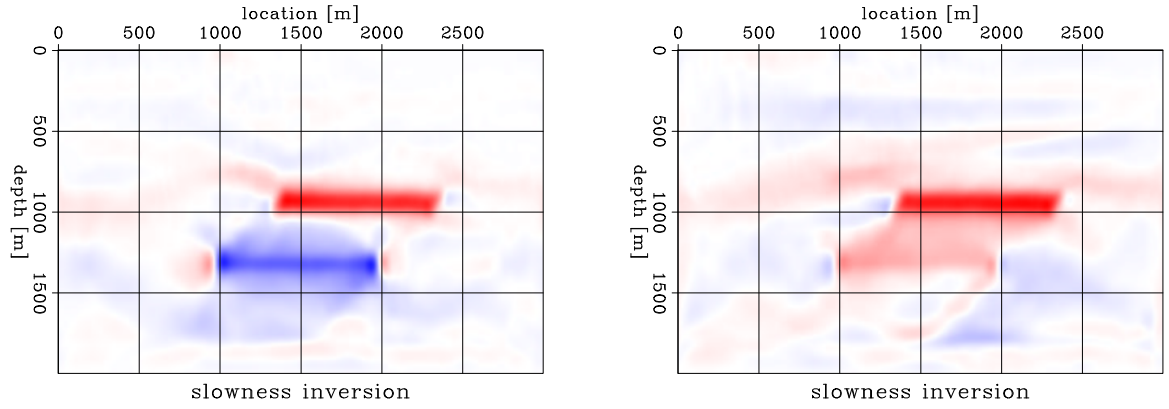


Figure 6: Slowness perturbation obtained using the least-squares inverse of the operator **L** in Equation (1). paul1-islow [CR]

only to changes of the physical parameters of the reservoir. Furthermore, since our processing is purely acoustic, we also need to insure that the image changes are dominated by changes of compressional slowness, and not by other elastic effects.

ACKNOWLEDGMENT

This work was performed during the first author's internship at BP Upstream Technology Group.

REFERENCES

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

Biondi, B., Deutsch, C., Mavko, G., Mukerji, T., Rickett, J., Thiele, M., Gundesø, R., and Lumley, D., 1996, Reservoir monitoring: A multidisciplinary feasibility study: 66th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1775–1778.

Kristiansen, P., Christie, P., Bouska, J., O'Donovan, A., Westwater, P., and Thorogood, E., 2000, Foinaven 4-D: Processing and analysis of two designer 4-Ds:, *in* 70th Ann. Internat. Mtg Soc. of Expl. Geophys., 1456–1459.

Lumley, D., 1995, Seismic time-lapse monitoring of subsurface fluid flow: Ph.D. thesis, Stanford University.

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

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 21–27

Short Note

Wave-equation MVA using diffracted data

Paul Sava and John Etgen[1]

INTRODUCTION

Migration velocity analysis (MVA) using diffracted data is not a new concept. Harlan (1986) addressed this problem and proposed a method to isolate diffraction events around faults. He also proposed a MVA technique applicable to simple geology, constant velocity or $v(z)$, and quantifies the focusing quality using statistical tools. de Vries and Berkhout (1984) use the concept of minimum entropy to evaluate diffraction focusing, and apply this methodology to MVA, again for the case of simple geology.

Biondi and Sava (1999) introduce a method of migration velocity analysis using wave-equation techniques (WEMVA), which aims at improving the quality of migrated images, mainly by correcting moveout inaccuracies of specular energy. The slowness model is estimated by finding the slowness perturbation which explains the difference between the migrated image using the reference model and an externally-defined target image (Sava and Symes, 2002).

The moveout information given by the specular energy is not the only information contained by an image migrated with the incorrect slowness. Non-specular diffracted energy is present in the image and clearly indicates slowness inaccuracies. Since a difference between an inaccurate image and a perfectly focused target image contains both specular and non-specular energy, WEMVA is naturally able to derive velocity updates based on both these types of information. In contrast, traveltime-based MVA methods cannot easily deal with the diffraction energy, and are most of the time concerned with moveout analysis.

In this paper, we examine the resolving power of the unfocused diffraction energy present in migrated images. The target applications of these techniques are in areas of complicated geology in which diffractions are abundant and can be clearly identified and isolated. Examples include highly fractured reservoirs, carbonate reservoirs, rough salt bodies and reservoirs with complicated stratigraphic features.

Of particular interest is the case of salt bodies. Diffractions can help estimate more accurate velocities at top of salt, particularly in the cases of rough salt bodies. Moreover, diffraction

[1]**email:** paul@sep.stanford.edu, john.etgen@bp.com

energy may be the most sensitive velocity information we have under salt, since most of the reflected energy we record at the surface has only a narrow range of angles of incidence at the reflector, rendering the analysis of moveout useless.

METHODOLOGY

In its original formulation (Biondi and Sava, 1999), wave-equation migration velocity analysis relates a perturbation of the slowness model ($\Delta\mathbf{S}$) to its corresponding perturbation of the seismic image ($\Delta\mathbf{R}$). Mathematically, this relation can be expressed as the linear fitting goal

$$\mathbf{L}\Delta\mathbf{S} \approx \Delta\mathbf{R}. \tag{1}$$

$\mathbf{L}$ is the WEMVA operator and is constructed as a linearization of downward continuation operators involving the Born approximation (Sava and Fomel, 2002). We obtain the slowness perturbation $\Delta\mathbf{S}$ from Equation (1) by applying either the adjoint or the least-squares inverse of $\mathbf{L}$ to the image perturbation $\Delta\mathbf{R}$.

The critical quantity in Equation (1) is the perturbation of the seismic image $\Delta\mathbf{R}$. For the purpose of this equation, this is the known quantity and various techniques can be used to derive it.

Suppose we can isolate all the diffractions from a given dataset. The migration velocity is correct if all diffractions are focused, both as a function of space and as a function of offset. Any inaccuracy of the velocity model leaves undiffracted energy in the image.

This simple observation gives us a mechanism to define image perturbations usable for WEMVA. First, we migrate the data using the reference slowness model and obtain a reference image $\mathbf{R}_0$. Second, we correct all unfocused diffractions using a residual technique (residual migration for example), and obtain an improved image $\mathbf{R}$. Finally we take the difference between $\mathbf{R}$ and $\mathbf{R}_0$ as the image perturbation $\Delta\mathbf{R}$, which can be inverted for slowness perturbation $\Delta\mathbf{S}$ using Equation (1).

This method is in principle usable for diffraction data analyzed in a prestack volume. However, as we will demonstrate with the example in the next section, a substantial part of information usable for MVA is present in the zero offset section. Our example concerns WEMVA purely at zero offset with the goal of isolating focusing effects from moveout-based effects.

EXAMPLE

We illustrate this diffraction focusing technique with a synthetic model simulating the diffracting points at the top of a rough salt body. The model is depicted in Figure 1: the reflectivity at the top, and the reference slowness in the middle. The model consists of several diffractions, and the reference slowness is smoothly spatially varying.

Figure 1 shows at the bottom the image perturbation $\Delta\mathbf{R}$ caused by the ideal slowness perturbation $\Delta\mathbf{S}$ shown in the top panel in Figure 2. $\Delta\mathbf{R}$ is created by subtracting the reference image from the perfectly focused one $\Delta\mathbf{R} = \mathbf{R} - \mathbf{R}_0$.

We take the image perturbation shown in the bottom panel of Figure 1 and compute the corresponding slowness perturbation using Equation (1). Figure 2 shows in the middle the result we obtain by applying the adjoint of the operator $\mathbf{L}$ to the image perturbation in Figure 1, and at the bottom the result of applying the least-squares inverse of $\mathbf{L}$ to the same $\Delta\mathbf{R}$.

Despite the inherent vertical smearing caused by the limited angular coverage, the slowness perturbations are nicely focused at their correct locations. Obviously, the result obtained with the least-squares inverse is much better focused than the one obtained by the simple adjoint operator, although we have only used the zero-offset and not the entire prestack data. The simple backprojection (top panel in Figure 2) creates "fat rays," also discussed by Woodward (1992) and Sava (2000).

CONCLUSIONS

We present an application of the WEMVA methodology to diffracted seismic data. Diffractions carry a substantial amount of velocity information which is largely ignored by the current MVA techniques. We show that diffractions can be used for accurate migration velocity analysis using the WEMVA methodology.

In the current form, the limitations of this method are identical to those of WEMVA, and are mainly related to the Born approximation assumed by the method which requires small perturbations. Further extensions are possible, but they remain subject to future research.

ACKNOWLEDGMENT

This work was performed during the first author's internship at BP Upstream Technology Group.

REFERENCES

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

de Vries, D., and Berkhout, A. J., 1984, Velocity analysis based on minimum entropy: Geophysics, **49**, no. 12, 2132–2142.

Harlan, W. S., 1986, Signal-noise separation and seismic inversion: Ph.D. thesis, Stanford University.

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

Sava, P., and Symes, W. W., 2002, A generalization of wave-equation migration velocity analysis: SEP–**112**, 27–36.

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

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, no. 1, 15–26.

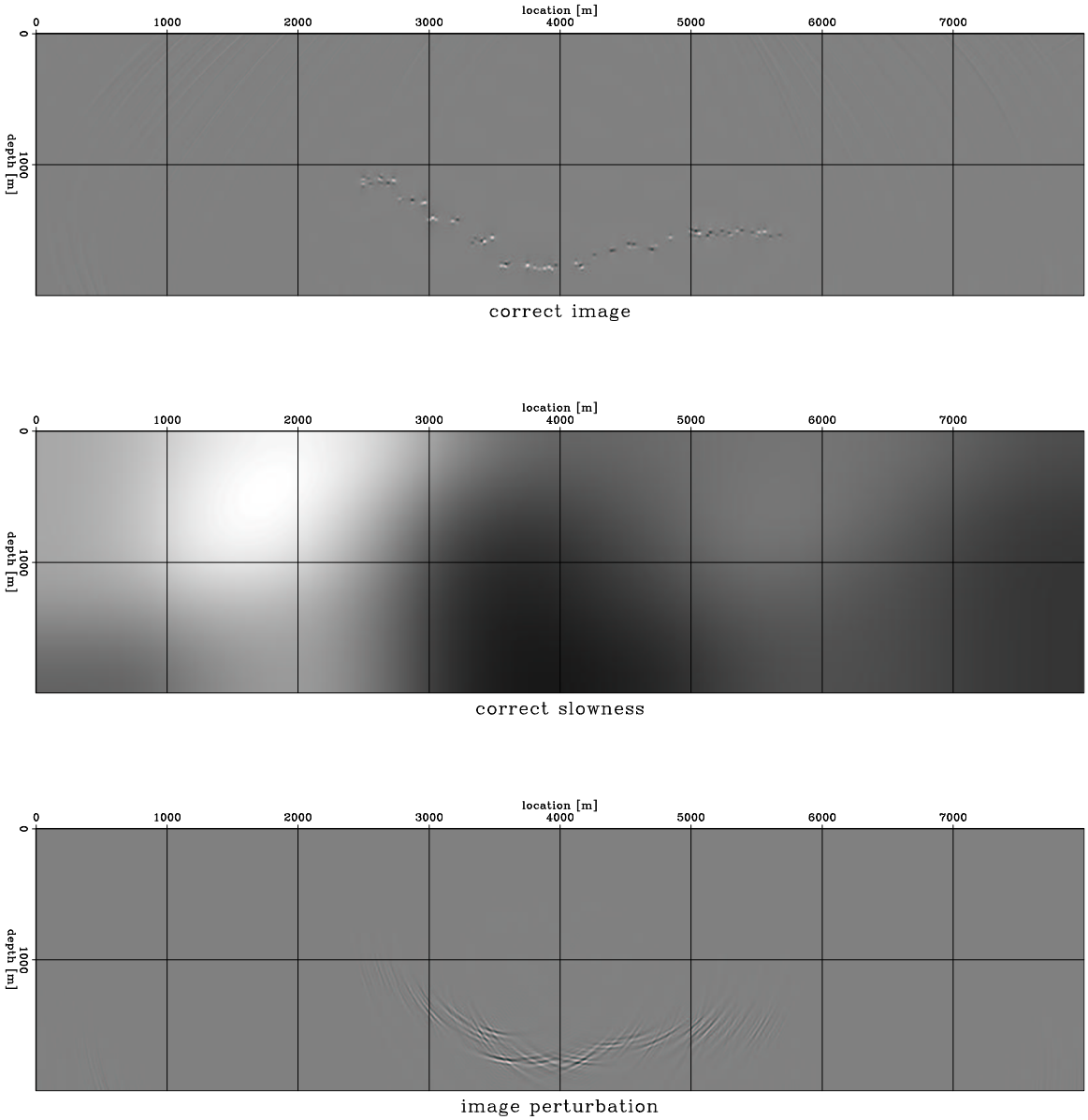


Figure 1: Synthetic model: reflectivity (top), background slowness (middle), and image perturbation (bottom). paul2-imags [CR]

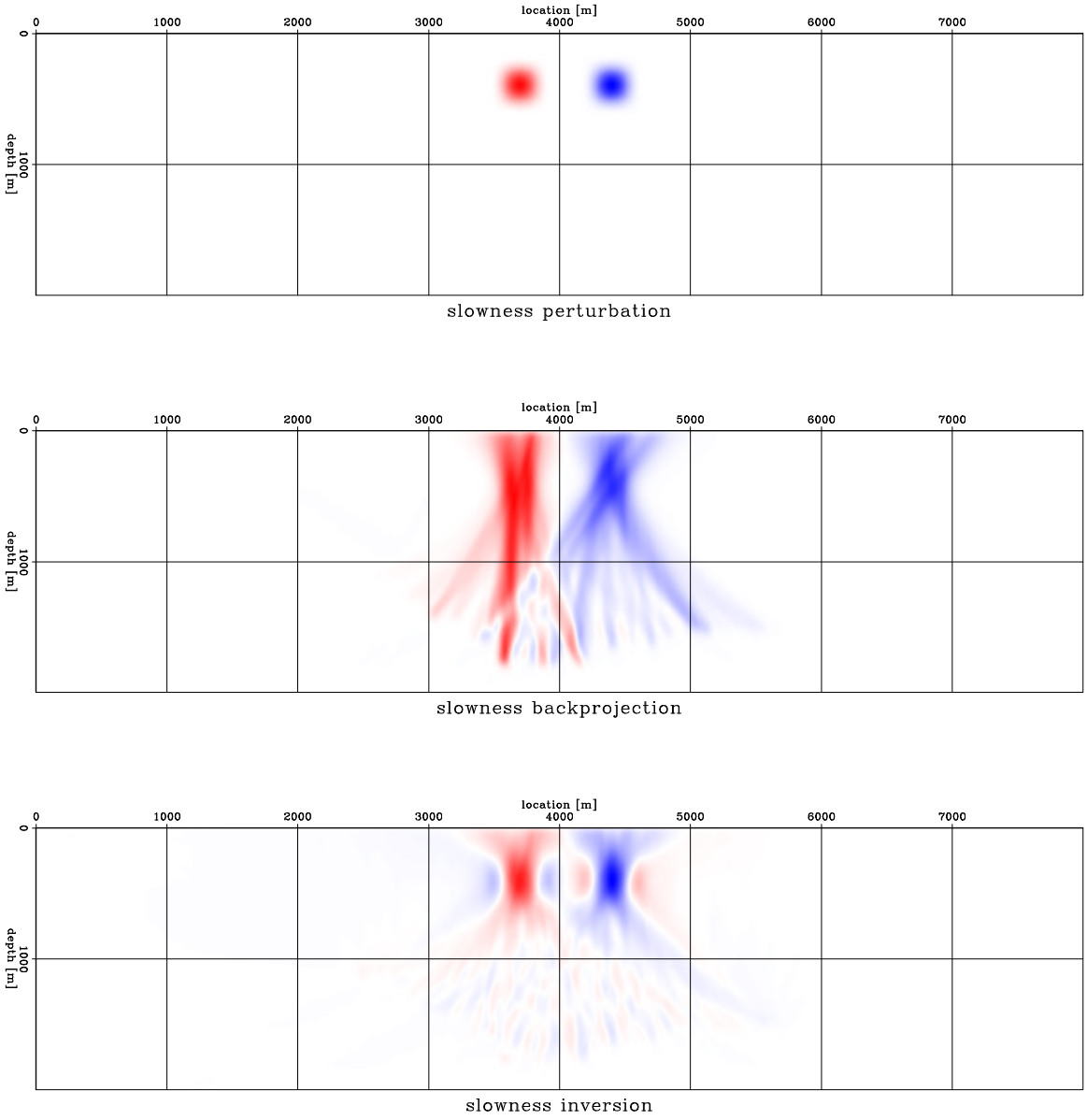


Figure 2: Ideal slowness perturbation (top), Slowness perturbation obtained by the adjoint of **L** (middle) and by the least-squares inverse of **L** (bottom) applied to the image perturbation in Figure 1. paul2-slows [CR]


stanford

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 27–37

A generalization of wave-equation migration velocity analysis

Paul Sava and William W. Symes[1]

ABSTRACT

Wave-equation migration velocity analysis is derived from wavefield-continuation migration techniques. The velocity model is updated by optimizing certain properties of the migrated images. Different migration velocity analysis optimization criteria exist, of which two commonly used are fitting a target image and minimizing the differential semblance of migrated images. Both techniques are special cases of a general family of optimization functions. Fitting a target image is an attractive technique because we can guide the solution in the desired direction. However, we can only progress in small steps with the target image being kept within the Born approximation with respect to the reference image. Minimizing differential semblance is an attractive technique, too, because we are operating with small differences of nearby offsets which are likely within the Born approximation. However, this method is not directly guided toward the solution and aliasing or any remnants of coherent noise, like multiples or converted waves, can cause it to diverge.

INTRODUCTION

Migration velocity analysis (MVA) is one of the most important problems of seismic imaging (Claerbout, 1999), and yet it remains one without a conventional solution. Many techniques have been devoted to solving this problem and, generally speaking, they fall into two broad categories: methods which directly use traveltimes computed using the eikonal equation, and methods which use the entire recorded wavefields. The methods in the first category are usually known by the name of *traveltime tomography* (Stork, 1992; Clapp, 2001), while the methods in the second category are known by the names of *wave-equation tomography* (Tarantola, 1984; Woodward, 1992) or *wave-equation migration velocity analysis* (Biondi and Sava, 1999; Sava and Fomel, 2002; Stolk and Symes, 2002).

The wave-equation MVA techniques are, in theory, superior to the traveltime-based MVA methods since they make use of the entire recorded data and not only of picked traveltimes at selected events. Some of those methods are also better able to account for multipathing occurring in complicated geological situations, a goal that is difficult to achieve with ray-traced traveltimes. Moreover, wave-equation techniques are more accurately describing wave propagation, since they are not based on high frequency asymptotic assumptions.

[1]**email:** paul@sep.stanford.edu, symes@caam.rice.edu

However, none of the wave-equation velocity analysis methods has yet been accepted as a practical solution to exploration problems. Part of the reason is cost, which remains high, despite the continually decreasing cost of computing hardware. In addition, many of those velocity analysis techniques become unstable if the data are polluted with coherent noise, if the recorded offsets are too short or if enough low frequencies are not available in the band of the data (Pratt, 1999).

Wave-equation MVA (WEMVA) is different from wave-equation tomography (WET) with respect to the domain in which each one computes residuals: WET operates in the data space, and estimates velocity by fitting the recorded data, while WEMVA operates in the image space, and estimates velocity by improving the quality of the migrated images. As for the traveltime tomography methods, estimating velocity in the migrated image space is a much more robust approach and more likely to converge to geologically meaningful solutions.

The usual property used for optimization is that of flat events measured along angle-domain common-image gathers. Optimal flatness in the angle-domain is equivalent to optimal focusing at zero-offset (Stolk and Symes, 2002), therefore explicit conversion to the angle-domain is not necessary. Similarly, we could use focusing along the spatial axes as well as focusing along offset in order to estimate migration velocity (Sava and Etgen, 2002)

In this paper, we generalize the wave-equation migration velocity analysis technique to include both the target image fitting method of Biondi and Sava (1999) and the differential semblance optimization method of Stolk and Symes (2002) in a unified framework. We show that both methods are just special cases of a more general technique. We discuss these two members of this general class of problems, and we point out that other more or less optimal methods exist.

In the following sections, we present in detail our generalization of the WEMVA method, followed by an example and a brief discussion of the results.

THEORY OF WAVE-EQUATION MVA

This section presents in detail our generalization of WEMVA. Throughout this section, we use the following notation conventions: we write $\mathbf{A}[x]$ when we mean $\mathbf{A}$ operates on x, and $\langle a,b\rangle$ when we mean the inner product of the vectors a and b.

We begin with a brief review of wavefield extrapolation, followed by a discussion of image transformations to the optimization domain, the objective functions and their gradient with respect to velocity. In the end, we restate our main results in the familiar SEP notation using linear fitting goals.

Imaging by wavefield extrapolation

Imaging by wavefield extrapolation (WE) is based on recursive continuation of the wavefields $\mathcal{U}$ from a given depth level to the next by means of an extrapolation operator $\mathbf{E}$:

$$\mathcal{U}_{z+\Delta z} = \mathbf{E}_z \left[\mathcal{U}_z \right]. \tag{1}$$

This recursive relation can also be explicitly written in matrix form as

$$\begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -\mathbf{E}_0 & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{E}_1 & \mathbf{1} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{E}_{n-1} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \mathcal{U}_0 \\ \mathcal{U}_1 \\ \mathcal{U}_2 \\ \vdots \\ \mathcal{U}_n \end{pmatrix} = \begin{pmatrix} \mathcal{D}_0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

or in a more compact notation as:

$$(\mathbf{1} - \mathbf{E})\, \mathcal{U} = \mathcal{D}, \tag{2}$$

where the vector $\mathcal{D}$ stands for the recorded data, $\mathcal{U}$ for the extrapolated wavefield, $\mathbf{E}$ for the extrapolation operator and $\mathbf{1}$ for the identity operator.

The wavefield at every depth level $\mathcal{U}_z$ is imaged using an imaging operator $\mathbf{I}_z$:

$$\mathcal{R}_z = \mathbf{I}_z \left[\mathcal{U}_z \right], \tag{3}$$

where $\mathcal{R}_z$ stands for the image at some depth level. We can write the same relation in compact matrix form as:

$$\mathcal{R} = \mathbf{I}\mathcal{U}, \tag{4}$$

where $\mathcal{R}$ stands for the image, and $\mathbf{I}$ stands for the imaging operator which is applied to the extrapolated wavefield $\mathcal{U}$.

Wavefield perturbations

A perturbation of the wavefield at some depth level can be derived from the background wavefield by a simple application of the chain rule to Equation (1):

$$\delta\mathcal{U}_{z+\Delta z} = \mathbf{E}_z \left[\delta\mathcal{U}_z \right] + \delta\mathbf{E}_z \left[\mathcal{U}_z \right]. \tag{5}$$

This is also a recursive equation which can be written in matrix form as

$$\begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -\mathbf{E}_0 & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{E}_1 & \mathbf{1} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{E}_{n-1} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \delta\mathcal{U}_0 \\ \delta\mathcal{U}_1 \\ \delta\mathcal{U}_2 \\ \vdots \\ \delta\mathcal{U}_n \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \delta\mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \delta\mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \delta\mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathcal{U}_0 \\ \mathcal{U}_1 \\ \mathcal{U}_2 \\ \vdots \\ \mathcal{U}_n \end{pmatrix},$$

or in a more compact notation as:

$$(\mathbf{1} - \mathbf{E})\, \delta\mathcal{U} = \delta\mathbf{E}\mathcal{U}, \tag{6}$$

where the operator $\delta\mathbf{E}$ stands for a perturbation of the extrapolation operator $\mathbf{E}$.

Biondi and Sava (1999) show that, at every depth level, we can write the operator $\delta\mathbf{E}$ as a chain of the extrapolation operator $\mathbf{E}$ and a scattering operator $\mathbf{S}$ applied to the slowness perturbation δs_z:

$$\delta\mathbf{E}_z\left[\mathcal{U}_z\right] = \mathbf{E}_z\left[\mathbf{S}_z\left[\delta s_z\right]\right]. \tag{7}$$

The expression for the wavefield perturbation $\delta\mathcal{U}$ becomes

$$\delta\mathcal{U}_{z+\Delta z} = \mathbf{E}_z\left[\delta\mathcal{U}_z\right] + \mathbf{E}_z\left[\mathbf{S}_z\left[\delta s_z\right]\right], \tag{8}$$

which is also a recursive relation that can be written in matrix form as

$$\begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ -\mathbf{E}_0 & \mathbf{1} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{E}_1 & \mathbf{1} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{E}_{n-1} & \mathbf{1} \end{pmatrix} \begin{pmatrix} \delta\mathcal{U}_0 \\ \delta\mathcal{U}_1 \\ \delta\mathcal{U}_2 \\ \vdots \\ \delta\mathcal{U}_n \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{E}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_1 & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{E}_{n-1} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{S}_0 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{S}_n \end{pmatrix} \begin{pmatrix} \delta s_0 \\ \delta s_1 \\ \delta s_2 \\ \vdots \\ \delta s_n \end{pmatrix},$$

or in a more compact notation as:

$$(\mathbf{1} - \mathbf{E})\,\delta\mathcal{U} = \mathbf{E}\mathbf{S}\delta s. \tag{9}$$

The vector δs stands for the slowness perturbation.

If we introduce the notation

$$\mathbf{G} = (\mathbf{1} - \mathbf{E})^{-1}\,\mathbf{E}\mathbf{S}, \tag{10}$$

we obtain a relation between a slowness perturbation and the corresponding wavefield perturbation:

$$\delta\mathcal{U} = \mathbf{G}\delta s. \tag{11}$$

Image transformation

Migration velocity analysis is based on estimating the velocity that optimizes certain properties of the migrated images. In general, measuring such properties involves making a transformation to the extrapolated wavefield by some function f, followed by imaging:

$$\mathcal{P}_z = \mathbf{I}_z\left[f_z(\mathcal{U}_z)\right]. \tag{12}$$

In compact matrix form, we can write this relation as:

$$\mathcal{P} = \mathbf{I}f(\mathcal{U}). \tag{13}$$

The image $\mathcal{P}$ is subject to optimization from which we derive the velocity updates.

Examples of transformation functions are:

- $f(x) = x - t$ where t is a known target. A WEMVA method based on this criterion optimizes

  $$\mathcal{P}_z := \mathbf{I}_z\left[\mathcal{U}_z - \mathcal{T}_z\right], \tag{14}$$

  where $\mathcal{T}_z$ stands for the target wavefield. For this method, we can use the acronym TIF standing for *target image fitting* (Biondi and Sava, 1999; Sava and Fomel, 2002).

- $f(x) = Dx$ where D is a known operator. A WEMVA method based on this criterion optimizes
$$\mathcal{P}_z := \mathbf{I}_z \left[\mathbf{D}_z \left[\mathcal{U}_z \right] \right]. \tag{15}$$
If **D** is a differential semblance operator, we can use the acronym DSO standing for *differential semblance optimization* (Symes and Carazzone, 1991; Stolk and Symes, 2002).

In general, both examples presented above belong to a family of affine functions that can be written as
$$\mathcal{P}_z := \mathbf{I}_z \left[\mathbf{A}_z \left[\mathcal{U}_z \right] - \mathbf{B}_z \left[\mathcal{T}_z \right] \right], \tag{16}$$
or in compact matrix form as
$$\mathcal{P} := \mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}), \tag{17}$$
where the operators **A** and **B** are known and take special forms depending on the optimization criterion we use. For example, $\mathbf{A} = \mathbf{1}$ and $\mathbf{B} = \mathbf{1}$ for TIF, and $\mathbf{A} = \mathbf{D}$ and $\mathbf{B} = \mathbf{0}$ for DSO. **1** stands for the identity operator, and **0** stands for the null operator.

Objective function

With the definition in Equation (17), we can write the optimization function J as:
$$J(s) := \frac{1}{2} \sum_{z,\vec{m},\vec{h}} |\mathcal{P}_z|^2 = \frac{1}{2} \sum_{z,\vec{m},\vec{h}} \left| \mathbf{I}_z \left[\mathbf{A}_z \left[\mathcal{U}_z \right] - \mathbf{B}_z \left[\mathcal{T}_z \right] \right] \right|^2, \tag{18}$$
where s is the slowness function, and $z, \vec{m}, \vec{h}$ stand respectively for depth, and the midpoint and offset vectors. In compact matrix form, we can write the objective function as:
$$J(s) := \frac{1}{2} |\mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T})|^2, \tag{19}$$
which takes special forms depending on our choice of the operators **A** and **B**:

WEMVA by TIF	WEMVA by DSO
$J(s) = \frac{1}{2} \left| \mathbf{I}(\mathcal{U} - \mathcal{T}) \right|^2$	$J(s) = \frac{1}{2} \left| \mathbf{I}(\mathbf{D}\mathcal{U}) \right|^2$

Gradient

Optimization of the objective function in Equation (19) requires computation of its gradient with respect to slowness. The objective function J can be rewritten using the inner product as:
$$J(s) = \frac{1}{2} \langle \mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}), \mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}) \rangle. \tag{20}$$
A perturbation of the function J is related to a perturbation of the wavefield by the relation:
$$\delta J(s) = \langle \mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}), \mathbf{I}\mathbf{A}\delta\mathcal{U} \rangle. \tag{21}$$

If we replace $\delta\mathcal{U}$ from Equation (11) we obtain:

$$\delta J(s) = \langle \mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}), \mathbf{IAG}\delta s \rangle, \tag{22}$$

therefore the gradient of the objective function can be written as

$$\nabla_s J = \mathbf{G}^*\mathbf{A}^*\mathbf{I}^*\mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}). \tag{23}$$

Following the definition of the operator $\mathbf{G}$, we can write

$$\mathbf{G}^* = \mathbf{S}^*\mathbf{E}^*\left[(\mathbf{1}-\mathbf{E})^{-1}\right]^* = \mathbf{S}^*\mathbf{E}^*\left[(\mathbf{1}-\mathbf{E})^*\right]^{-1}. \tag{24}$$

Finally, the expression for the gradient of the objective function with respect to slowness becomes

$$\nabla_s J = \mathbf{S}^*\mathbf{E}^*\left[(\mathbf{1}-\mathbf{E})^*\right]^{-1}\mathbf{A}^*\mathbf{I}^*\mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}) \tag{25}$$

which takes special forms depending on our choice of the operators $\mathbf{A}$ and $\mathbf{B}$:

WEMVA by TIF	WEMVA by DSO
$\nabla_s J = \mathbf{S}^*\mathbf{E}^*\left[(\mathbf{1}-\mathbf{E})^*\right]^{-1}\mathbf{I}^*\mathbf{I}(\mathcal{U}-\mathcal{T})$	$\nabla_s J = \mathbf{S}^*\mathbf{E}^*\left[(\mathbf{1}-\mathbf{E})^*\right]^{-1}\mathbf{D}^*\mathbf{I}^*\mathbf{ID}\mathcal{U}$

The gradient in Equation (25) is computed using the adjoint state method, which can be summarized by the following steps:

1. Compute by downward continuation the wavefield

$$\mathbf{A}^*\mathbf{I}^*\mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}). \tag{26}$$

2. Compute by upward continuation the adjoint state wavefield

$$\mathcal{W} = \left[(\mathbf{1}-\mathbf{E})^*\right]^{-1}\mathbf{A}^*\mathbf{I}^*\mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}), \tag{27}$$

   i.e. solve the adjoint state system

$$(\mathbf{1}-\mathbf{E})^*\,\mathcal{W} = \mathbf{A}^*\mathbf{I}^*\mathbf{I}(\mathbf{A}\mathcal{U} - \mathbf{B}\mathcal{T}). \tag{28}$$

3. Compute the gradient

$$\nabla_s J = \mathbf{S}^*\mathbf{E}^*\mathcal{W}. \tag{29}$$

Linearization

Minimizing the objective function J in Equation (19) involves solving a non-linear least-squares problem using the gradient given by Equation (25).

Alternatively, we can linearize the wavefield $\mathcal{U}$ with respect to a reference wavefield $\mathcal{U}^r$

$$\mathcal{U} = \mathcal{U}^r + \delta\mathcal{U} = \mathcal{U}^r + \mathbf{G}\delta s \tag{30}$$

and optimize

$$J(s) = \frac{1}{2}\left|\mathbf{I}\left(\mathbf{A}\mathcal{U}^r - \mathbf{B}\mathcal{T} + \mathbf{A}\mathbf{G}\delta s\right)\right|^2, \tag{31}$$

which is a linear least-squares problem.

Equation (31) can also be represented by fitting goals using the usual SEP terminology as:

$$-\mathbf{I}\left(\mathbf{A}\mathcal{U}^r - \mathbf{B}\mathcal{T}\right) \approx \mathbf{I}\mathbf{A}\mathbf{G}\delta s, \tag{32}$$

which takes special forms depending on our choice of the operators **A** and **B**:

WEMVA by TIF	WEMVA by DSO
$-\mathbf{I}(\mathcal{U}^r - \mathcal{T}) \approx \mathbf{I}\mathbf{G}\delta s$	$-\mathbf{I}(\mathbf{D}\mathcal{U}^r) \approx \mathbf{I}\mathbf{D}\mathbf{G}\delta s$

EXAMPLE

Our synthetic example, Figure 1, is represented by a simple horizontal reflector (top panel) embedded in a velocity model with smooth lateral velocity variation (middle panel). A small perturbation introduced in the velocity model creates an image perturbation which does not violate the Born approximation (bottom panel).

We compute the gradient of the objective function in Equation (19) using Equation (25) particularized both for the target image fitting (TIF) and for differential semblance optimization (DSO) criteria. Figure 2 shows the ideal image perturbation (top panel), the gradient for TIF (middle panel), and the gradient for DSO (bottom panel).

DISCUSSION

Not surprisingly, the two methods generally represented by Equation (17) produce significantly different results. Although our analysis in this paper does not cover all cases and possibilities, we can make several observations:

- The general form in Equation (17) is not unique, meaning that other forms of equal generality exist. Moreover, we have presented and compared just two members of our general form, although many others exist. The obvious question, for which we do not have a definite answer, is which is the optimal form? Is there such thing, or do we need to consider different forms for different situations? These questions remain the subject of future research.

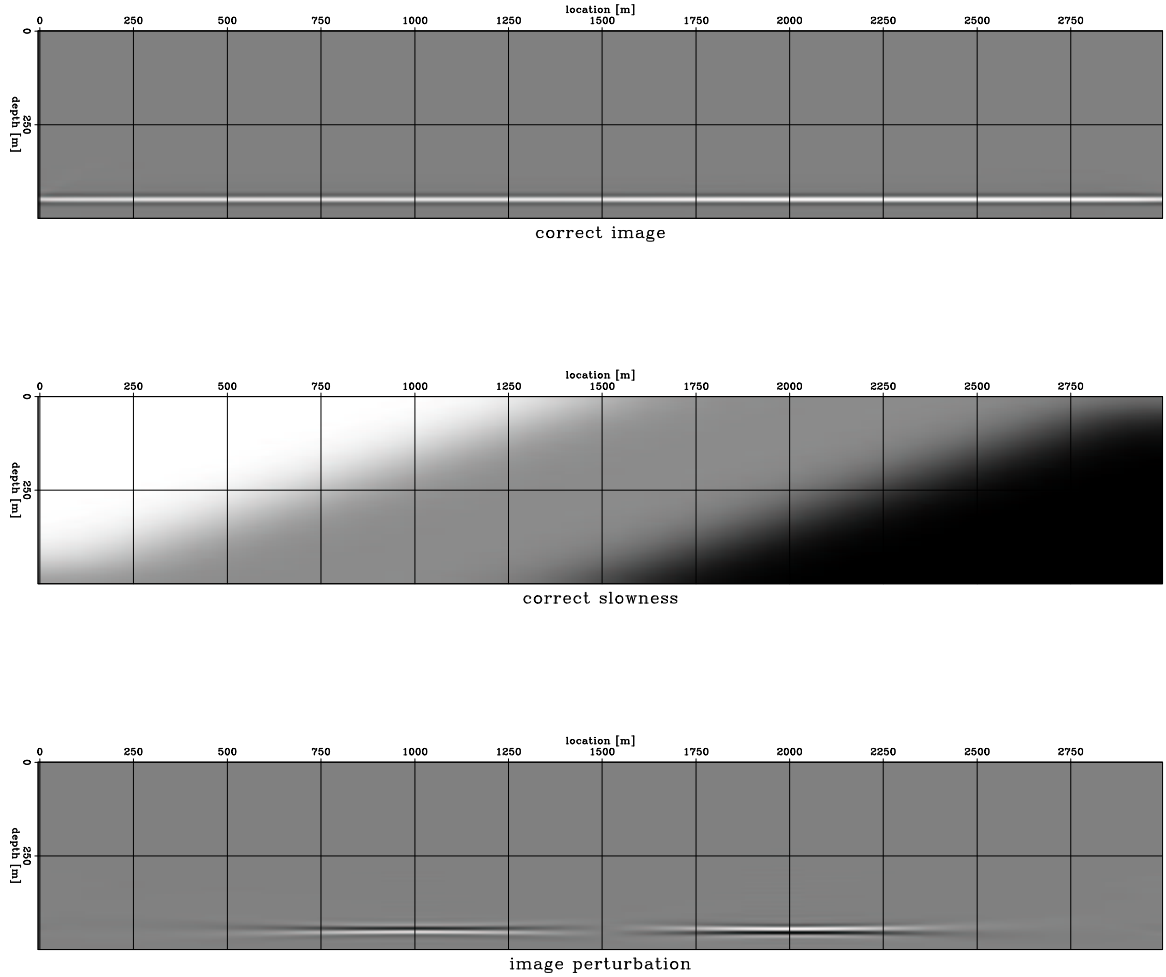


Figure 1: Reflectivity model (top), background slowness (middle), and image perturbation (bottom). paul3-imags [CR]

- The *target image fitting* (TIF) approach is an attractive alternative because it can, in principle, be driven in the desired direction given by the target image. However, if the constraints presented by the Born approximation are not observed (i.e. the target is too far from the actual image), then inversion may diverge (Sava and Fomel, 2002). We also need to create the actual target, an improved image, which is not a trivial task.

- The *differential semblance optimization* (DSO) approach is also attractive for other reasons. The objective function is smooth and unimodal, at least for certain simplified cases (Symes, 1999). However, even DSO is not guaranteed to converge when the data are aliased or when they are polluted with residual multiples or converted waves.

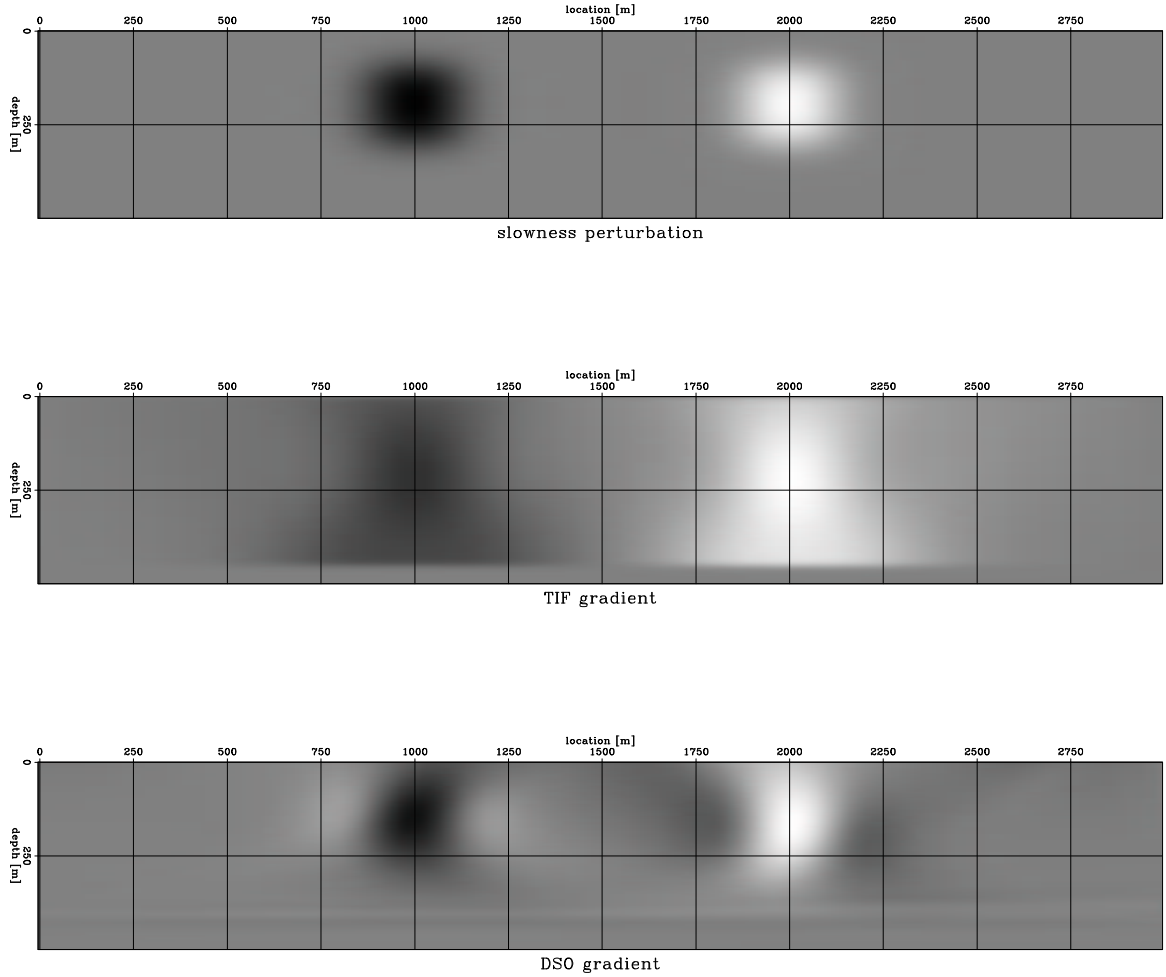


Figure 2: Slowness perturbation (top), TIF gradient (middle) and DSO gradient (bottom). paul3-slows [CR]

CONCLUSION

We present a generalization of the wave-equation migration velocity analysis technique. We show that various objective functions can be used and that each one of them has distinctive properties which make them attractive under different circumstances. Our generalization provides a framework in which we can test various optimization strategies. Those extensions, however, remain subjects for future research.

ACKNOWLEDGMENT

We thank Biondo Biondi for inspiring discussions.

REFERENCES

Biondi, B., and Sava, P., 1999, Wave-equation migration velocity analysis: 69th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1723–1726.

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.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model: Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model:, Soc. of Expl. Geophys., Geophysics, 888–901.

Sava, P., and Etgen, J., 2002, Wave-equation mva using diffracted data: SEP–**112**, 21–26.

Sava, P., and Fomel, S., 2002, Wave-equation migration velocity analysis beyond the Born approximation: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Stolk, C. C., and Symes, W. W., 2002, Theory of differential semblance velocity analysis by wave-equation migration: The Rice Inversion Project, ftp://ftp.trip.caam.rice.edu/private/trip/AR00_ps/wemvaseg.ps.

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

Symes, W. W., and Carazzone, J. J., 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, no. 5, 654–663.

Symes, W. W., 1999, All stationary points of differential semblance are asymptotic global minimizers: Layered acoustics: SEP–**100**, 71–92.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, no. 8, 1259–1266.

Woodward, M. J., 1992, Wave-equation tomography: Geophysics, **57**, no. 1, 15–26.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 37–47

Dealing with errors in automatic velocity analysis

Robert G. Clapp[1]

ABSTRACT

The lack of human interaction in automatic reflection tomography leads to a larger percentage of "bad" data points. The number of data points associated with events with spurious moveouts (such as multiples and converted waves) can be minimized by intelligently controlling the semblance scanning range. The effect of the bad data points can be limited by replacing the standard L_2 norm solution with a norm closer to L_1 by reweighted least-squares. By replacing the standard constant ϵ parameter with a diagonal operator, areas with large errors in moveout can be highly regularized with minimal effect on areas with more reliable moveout information. This methodology is applied to a complex 2-D dataset.

INTRODUCTION

In ray-based reflection tomography, picking reflectors is an integral and painful part of the process (Clapp, 2001b; van Trier, 1990; Stork, 1992; Kosloff et al., 1996). The common methodology is to pick a series of reflectors from a migrated image. A set of rays are then calculated that reflect at the picked interfaces. A major problem is the human intensive nature of reflector picking, especially for 3-D data. Automatic pickers can help, but significant human quality control (QC) is still necessary. A high level of QCing is required because inaccurate reflector picks lead to inaccurate reflector dip estimates. These poor estimates cause information to be back projected to the wrong portion of the model space, seriously hampering the inversion.

Woodward et al. (1998) and Clapp (2001a) introduced methods to limit the amount of picking required by selecting back projection points based on criteria such as semblance and dip coherence. These methods are successful in reducing the human cost of tomography but have two significant weaknesses. First, moveout is often characterized by a single value. This value is obtained by scanning over a range of moveouts and then selecting the maximum. In areas with significant multiple or converted wave energy, they will often have trouble distinguishing primary reflections (signal) from multiple and converted wave reflections (noise). The second problem is that these automatic point selection methods are generally going to have a larger level of erroneous moveout descriptions that generally increase with depth. These erroneous moveouts will generally cause large residuals which can dominate the inversion procedure.

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

In this paper, I show three simple methods to combat both problems. Unreasonable moveouts can be avoided by scanning over a large range of moveouts but only selecting points whose maximum is in a narrower range of moveouts. Second, I show that we can account for a higher level of variance by adding a diagonal weight to our model styling goal. Finally, I show that the effect of the remaining non-primary events and other erroneous moveouts can be further diminished by using re-weighted least-squares (Claerbout, 1998) to simulate a L_1 inversion problem where noisy data points have less of an effect.

LIMITING THE SCANNING RANGE

In Clapp (2001a), I outlined a procedure for selecting points for back projection. The goal was to find points with high dip coherence and semblance at a minimum distance from each other. This methodology can run into problems for events whose moveout doesn't correspond to primary events or whose moveout is not adequately defined by calculating vertical semblance. For example, the common reflection point (CRP) gathers in Figure 1 shows every fifth gather along the left edge of a salt body. Note the coherent but "hockey stick" like shapes within the 'A' oval. These can be caused by small velocity errors (Biondi and Symes, 2002) but measuring just vertical moveout would indicate much larger errors. Clapp (2002) shows one way to address the latter concern.

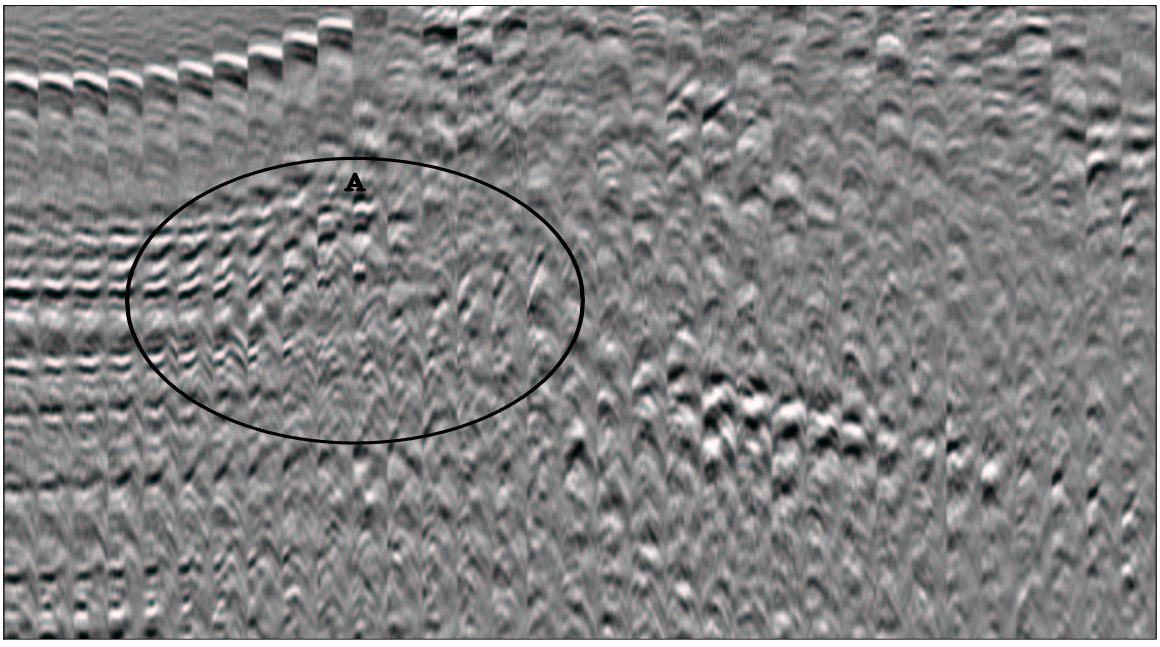


Figure 1: Every 5th gather to the left edge of a salt body. Note the coherent, "hockey stick" behavior within 'A'. bob3-gathers [CR]

Simply limiting the range of acceptable moveouts that we search isn't a sufficient solution because the maximum often will be at the extreme scan range. A simple methodology to minimize the effect of unreasonable moveouts is to scan over a large range of acceptable

moveouts and only accept points whose maximum fall within a smaller range (see Figure 2). With this methodology, spurious moveouts can be identified and ignored. When dealing with internal multiples or events whose moveout is close to acceptable, failure can still result.

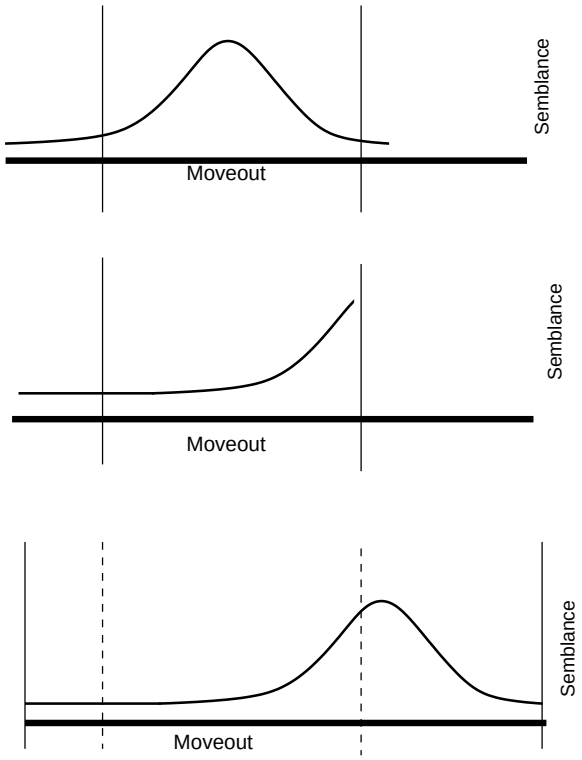


Figure 2: The top figure shows an example of a good point. The maximum is reasonable and within the scanning region indicated by the solid vertical lines. The second plot shows the problematic situation. The moveout is unreasonable and its maximum is outside the scanning range. We can avoid using the unrealistic moveout by scanning over moveouts between the solid lines but only selecting points whose maximum is within the dashed lines, bottom panel. bob3-limited [NR]

To show the benefits of this methodology, I applied it to a complex 2-D dataset. Figure 3 shows an initial velocity model and migrated image of a 2-D line from a 3-D dataset donated by Total Fina Elf. Figure 4 shows the updated velocity model and migrated image without limiting the scan range. Note the extreme velocity along the edge of the salt. The resulting image is less coherent than the initial image, especially in the ovals indicated by 'A', 'B', and 'C'. Figure 5 shows the result of limiting the range of acceptable moveouts. Note how the velocity along the edge is more reasonable. We see a strong salt bottom reflection at 'A', better definition of the valley at 'B', and more coherent events leading up to the salt edge at 'C'.

VARIABLE EPSILON

Let's begin with a review of the general fitting goals we are using for the tomography problem.

Tomography Review

Following the methodology of Clapp and Biondi (1999), I will begin by considering a regularized tomography problem. I will linearize around an initial slowness estimate and find a linear operator in the vertical traveltime domain $\mathbf{T}$ between our change in slowness $\Delta\mathbf{s}$ and

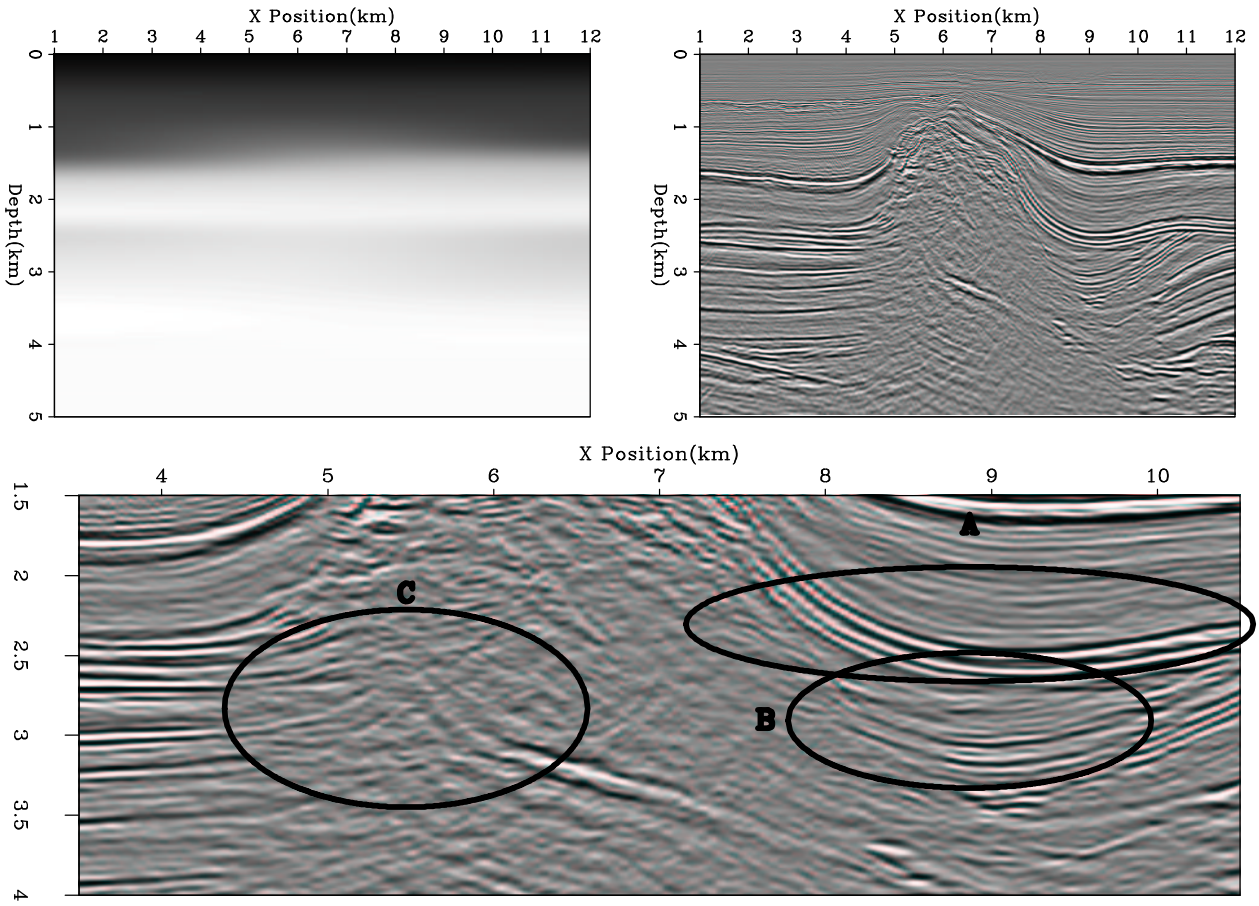


Figure 3: The starting model and migration of a 2-D line from a 3-D North Sea dataset. The top-left panel shows the velocity model (white indicates large velocities) and the top-right panel shows the migrated image using this velocity. The bottom panel shows a zoomed area around the salt body. Note the salt bottom,'A'; the valley structure at 'B'; and over the salt under-hang at 'C'. bob3-combo.vel0 [CR,M]

our change in traveltimes $\Delta\mathbf{t}$. We will write a set of fitting goals,

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\Delta\mathbf{s} \\ \mathbf{0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}, \end{aligned} \tag{1}$$

where $\mathbf{A}$ is our steering filter operator (Clapp et al., 1997) and ϵ is a Lagrange multiplier. However, these fitting goals don't accurately describe what we really want. Our steering filters are based on our desired slowness rather than change of slowness. With this fact in mind, we can rewrite our second fitting goal as:

$$\begin{aligned} \mathbf{0} &\approx \epsilon\mathbf{A}(\mathbf{s_0}+\Delta\mathbf{s}) \\ -\epsilon\mathbf{A}\mathbf{s_0} &\approx \epsilon\mathbf{A}\Delta\mathbf{s}. \end{aligned} \tag{2}$$

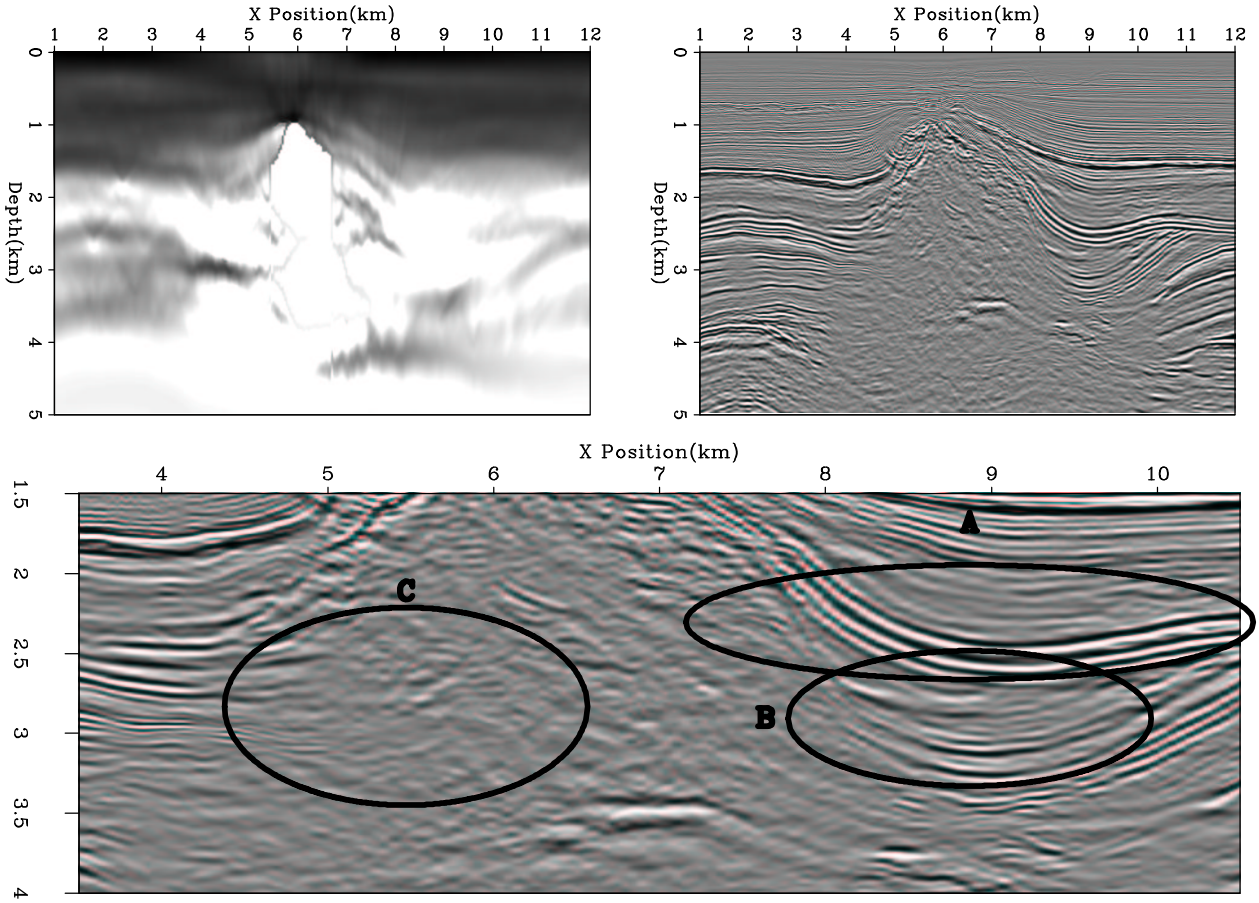


Figure 4: Data after one non-linear iteration. The top-left panel shows the velocity model and the top-right panel shows the migrated image using this velocity. The bottom panel shows a zoomed area around the salt body. Note the salt bottom,'A'; the valley structure at 'B'; and under the salt over-hang at 'C'. bob3-combo.vel1.bad [CR,M]

Our second fitting goal can not be strictly defined as regularization but we can still do a pre-conditioning substitution (Fomel et al., 1997), giving us a new set of fitting goals:

$$\begin{aligned} \Delta\mathbf{t} &\approx \mathbf{T}\mathbf{A}^{-1}\mathbf{p} \\ -\epsilon\mathbf{A}\mathbf{s_0} &\approx \epsilon\mathbf{I}\mathbf{p}. \end{aligned} \tag{3}$$

Our standard inversion fitting goals (3) make an assumption that our data fitting goal is equally believable everywhere. Stated another way, we want the same weight ϵ for our model styling goal everywhere. This is generally untrue. We can, and should, account for differing level of confidence in two different ways. If we have a measure of certainty about a data point (for example how much of a peak our semblance pick is) we can add a data covariance operator $\mathbf{W}$ to our fitting goals,

$$\begin{aligned} \mathbf{W}\Delta\mathbf{t} &\approx \mathbf{W}\mathbf{T}\mathbf{A}^{-1}\mathbf{p} \\ -\mathbf{A}\mathbf{s_0} &\approx \epsilon\mathbf{I}\mathbf{p}. \end{aligned} \tag{4}$$

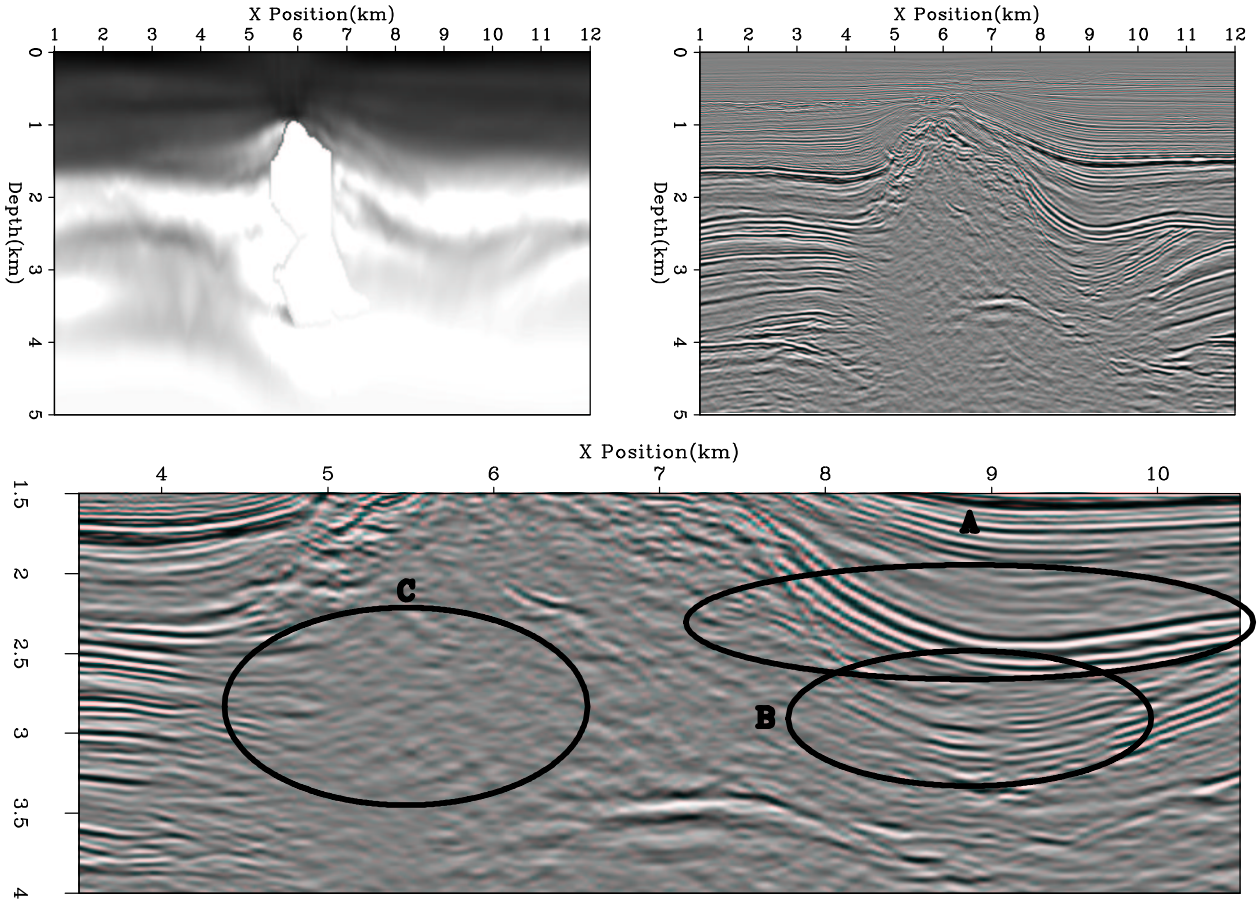


Figure 5: Data after one non-linear iteration with limited semblance search window. The top-left panel shows the velocity model and the top-right panel shows the migrated image using this velocity. The bottom panel shows a zoomed area around the salt body. Note the salt bottom,'A'; the valley structure at 'B'; and under the salt over-hang at 'C'. Note the improvements compared to Figure 4. bob3-combo.vel1.l2 [CR,M]

We can also often make statements about our confidence in our data fitting goal as a function of our model space. For example, generally as we get deeper, we will have less confidence in the points, and be less able to get a high frequency velocity model. We can account for this uncertainty by replacing the constant epsilon of fitting goal (4) with a diagonal weighting operator **E** resulting in the updated fitting goals,

$$\begin{aligned} \mathbf{W}\Delta\mathbf{t} &\approx \mathbf{W}\mathbf{T}\mathbf{A}^{-1}\mathbf{p} \\ -\mathbf{E}\mathbf{A}\mathbf{s_0} &\approx \mathbf{E}\mathbf{I}\mathbf{p}. \end{aligned} \tag{5}$$

By having this additional freedom we can allow for more model variability in the near surface and force more smoothing at deeper locations. Figure 6 shows the result of using the new fitting goals (5). Note how we have a higher frequency velocity structure above and a smoother below. The overall image quality is also improved compared to Figure 5.

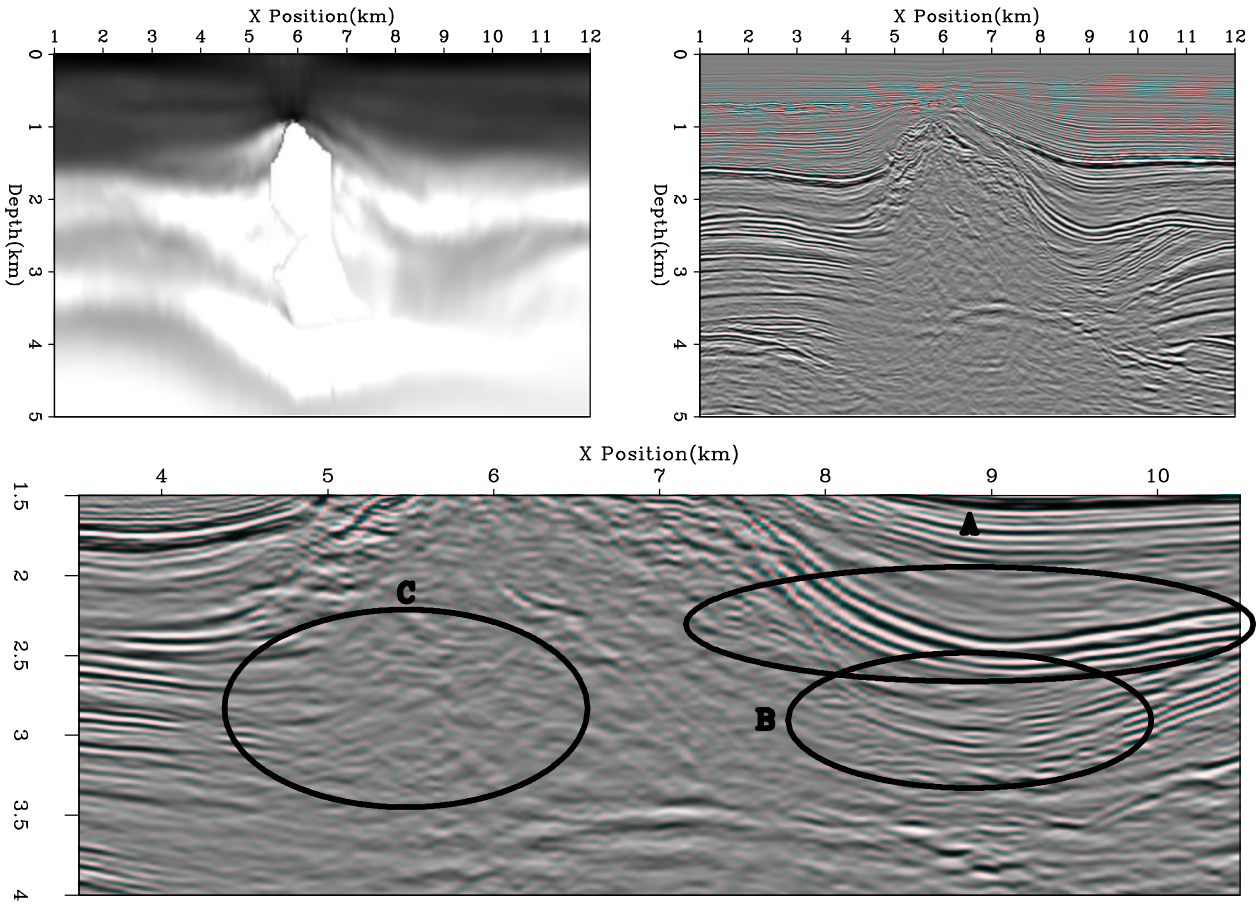


Figure 6: Data after one iteration using a constant ϵ. The top-left panel shows the velocity model and the top-right panel shows the migrated image using this velocity. The bottom panel shows a zoomed area around the salt body. Note the salt bottom,'A'; the valley structure at 'B'; and under the salt over-hang at 'C'. Note the improved image quality compared to Figure 4 and Figure 5. bob3-combo.vel1.eps [CR,M]

L_1 NORM TO HANDLE LARGE RESIDUALS

The previous section discussed a method to reduce the number of bad data points. Another approach is to limit their effect in the inversion. Generally we do not iterate to convergence. Early iterations tend to concentrate on large residuals. Erroneous data points tend to cause the large residuals. The result is that our solutions tend to be dominated by these erroneous data points.

A method to combat this problem is to change our misfit functional from the traditional

$$\mathbf{r_d} = \|\mathbf{d} - \mathbf{Lm}\|^2. \tag{6}$$

There are two different methods to change the misfit function. The first is to use a non-linear solver. With a non-linear solver there are a variety of misfit functions, most interestingly the

Ł$_1$ and Huber functionals (Huber, 1973; Claerbout, 1996; Clapp and Brown, 1999). A second approach is reweighted least-squares, (Nichols, 1994; Guitton, 2000). In reweighted least-squares a weighting operator is applied to the residuals based upon the size of residuals at certain points in the inversion.

The total non-linear approach has fewer parameters to manipulate and is generally more robust than reweighted approach. The downside of the completely non-linear approach is that it is significantly slower (a factor of ten or more is not uncommon). I chose the reweighted least-squares approach because I am most interested in finding and minimizing the effect of the largest residuals. I found that a single calculation of the weighting function after $\frac{1}{3}$ of the total number of iterations was sufficient to minimize the most troublesome residuals. Figure 7 shows the velocity and migration result using an L_1 norm. Note the improvement in image quality over either of the previous approaches (Figures 4 and 5). The salt bottom is more continuous, 'A'. The valley structure is better defined, 'B'. The reflectors are more continuous and extend closer to the salt at 'C'.

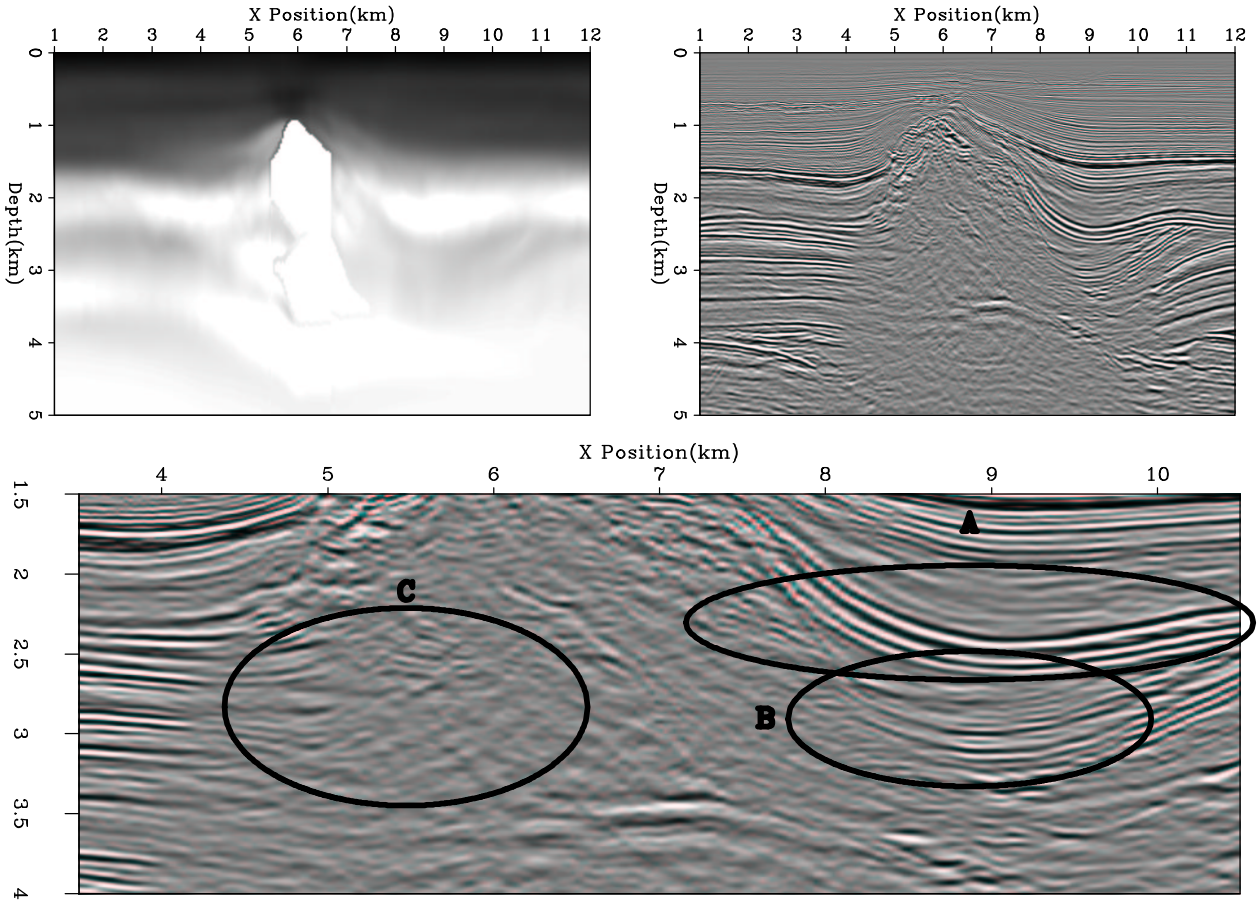


Figure 7: Data after one non-linear iteration using a reweighted least-squares. The top-left panel shows the velocity model and the top-right panel shows the migrated image using this velocity. The bottom panel shows a zoomed area around the salt body. Note the salt bottom,'A'; the valley structure at 'B'; and under the salt over-hang at 'C'. Note the improved image quality compared to Figure 4 and Figure 5. bob3-combo.vel1.steer [CR,M]

CONCLUSIONS

Automatic velocity updating has great potential for reducing processing cycle time. The problem of automatic methods selecting unreasonable moveouts can be reduced by scanning over a large moveout range and selecting events which fall within a smaller window. The effect of bad points can be further reduced by replacing the standard L_2 norm with a L_1 norm. Preliminary results are promising.

REFERENCES

Biondi, B., and Symes, W. W., 2002, Transformation to dip-dependent common image gathers: SEP–**112**, 65–82.

Claerbout, J., 1996, Conjugate-direction Huber regression: SEP–**92**, 229–235.

Claerbout, J. Geophysical Estimation by Example: Environmental soundings image enhancement:. http://sepwww.stanford.edu/sep/prof/, 1998.

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

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

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.

Clapp, R. G., 2002, Ray based tomography using residual Stolt migration: SEP–**112**, 1–14.

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

Guitton, A., 2000, Huber solver versus IRLS algorithm for quasi L1 inversion: SEP–**103**, 255–271.

Huber, P. J., 1973, Robust regression: Asymptotics, conjectures, and Monte Carlo: Ann. Statist., **1**, 799–821.

Kosloff, D., Sherwood, J., Koren, Z., MacHet, E., and Falkovitz, Y., 1996, Velocity and interface depth determination by tomography of depth migrated gathers: Geophysics, **61**, no. 5, 1511–1523.

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

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

van Trier, J., 1990, Tomographic determination of structural velocities from depth migrated seismic data: Ph.D. thesis, Stanford University.

Woodward, M. J., Farmer, P., Nichols, D., and Charles, S., 1998, Automated 3-D tomographic velocity analysis of residual moveout in prestack depth migrated common image point gathers:, *in* 68th Ann. Internat. Mtg Soc. of Expl. Geophys., 1218–1221.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 47–65

Velocity estimation for seismic data exhibiting focusing-effect AVO (part 2)

Ioan Vlad[1]

ABSTRACT

Vlad and Biondi (2002) have shown that focusing-effect AVO (FEAVO) exists and is visible in the angle domain. They have conjectured that wave-equation migration velocity analysis (WEMVA) might solve the FEAVO problem. I continue that line of work by showing evidence that WEMVA is the right tool, and by redoing the preprocessing and velocity analysis of the dataset on which FEAVO was defined.

INTRODUCTION

Small velocity anomalies can create AVO effects by focusing the reflected seismic wavefield. This impedes AVO analysis. The AVO anomalies caused by focusing are distinguishable by surface-consistent patterns. Vlad and Biondi (2002) named this phenomenon focusing-effect AVO (FEAVO). In this paper, I present and discuss the feasibility of a method whose application to FEAVO-affected 2D and 3D datasets would produce an estimate of the anomalies. The goal is to determine a velocity field accurate enough to generate a FEAVO-free prestack volume by downward continuing the wavefield through the FEAVO-generating anomalies. The accurate velocity model should be obtained by using the inversion method known as wave equation migration velocity analysis (WEMVA) (Biondi and Sava, 1999), with a fitting goal modified so that it is specifically geared toward extracting FEAVO effects.

I have continued that line of work by going into more depth with regard to each of the aspects of the study. I will respect the same structure. The material presented at each point consists of the progress in knowledge and research accomplished since the previous paper. I begin with a review of previous work used as a basis for my work, then I explain my approach and show evidence of its validity.

PREVIOUS WORK BY OTHERS

Beyond the mere "historical" interest, previous work on defining and describing the FEAVO effect can help answer some potential objections to the proposed approach. For example, one objection might be that FEAVO effects are randomly distributed. Therefore, their expression in

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

AVO pickings is the same as that of random noise, and can be treated the same way. White et al. (1988) showed that focusing and defocusing caused by a random distribution of small velocity anomalies will not lead to the canceling of the FEAVO effects by each other. Moreover, they prove that even if the distribution of velocity anomalies is random, the probability that FEAVO effects appear is not randomly distributed with the distance from the shot. Therefore, FEAVO presence cannot be equated with random noise in AVO picking, since FEAVO-causing anomalies are not randomly distributed in real data. In the case of shallow anomalies the distribution is influenced by depositional patterns [the "Rotten alligators" in Claerbout (1985)], and in the case of deep-origin FEAVO, by the presence of faults (Hatchell, 2000). Besides, the amplitude of the FEAVO effects is often so large and spatially extended that it renders AVO analysis impossible.

Attempts to invert FEAVO-affected data for a velocity model

The main papers discussing aspects of solving the FEAVO problem are Kjartansson (1979), Woodward (1987), Claerbout (1993), Bevc (1994), and Harlan (1994). All are iterative inversions that try to find the velocity model that will result in the given raw data. With the exception of Harlan (1994), this operator is some form of plane-wave decomposition using a straight-ray assumption; none uses an actual differential ("wave-equation") operator. Woodward (1987) applies corrections in order to account for differences between infinite-frequency rays and the "fat rays" associated with the physics of wave propagation. All of them, with the exception of Harlan (1994), invert either for the traveltime effects associated with AVO, or for the amplitudes, but not for both simultaneously.

None of the previous attempts is completely successful in producing a velocity model that satisfies the initial goals of the problem. Using only a two-dimensional midpoint-offset map of FEAVO effects, instead of recognizing that they correlate across depth (as equation 3 and Figure 9 show), they introduced too many degrees of freedom in the inversions. Considering only straight rays was incorrect even in the case of the universally encountered v(z) variation. All methods require some form of picking, which results in endless headaches. None makes successful use of the entire quantity of information by simultaneously considering both the traveltimes and the amplitudes for all the reflectors. Only one (Harlan, 1994) incorporates the information given by standard velocity analysis (which is not significantly affected by FEAVO effects).

Wave Equation Migration Velocity Analysis

Like tomography, migration velocity analysis (MVA) is a velocity analysis method by iterative inversion. The difference between the two is at a "strategic" level: instead of trying to find the velocity model that will result in the given raw data, it tries to find the velocity through which prestack migrated data results in a perfectly focused image. The chapter pertaining to velocity analysis in Biondi (2001) shows 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).

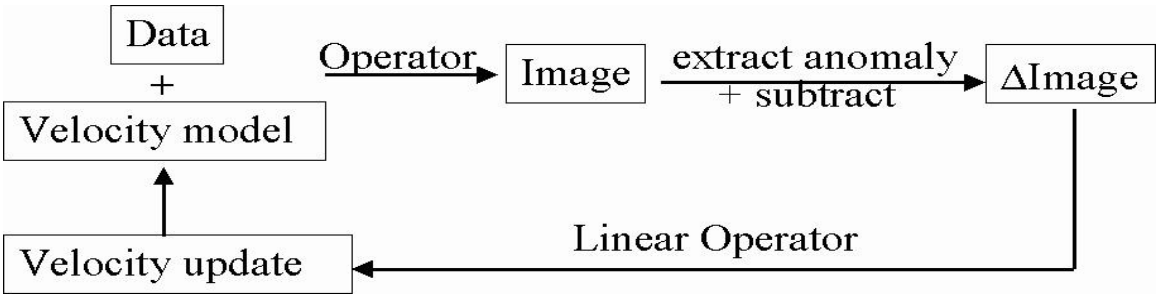


Figure 1: Migration Velocity Analysis flowchart nick2-invpic [NR]

MVA can be done with infinite-frequency (ray-based) operators (Clapp, 2001) or with finite band (wave-equation) operators (Biondi and Sava, 1999). The advantages of the latter over the former are detailed in the WEMVA chapter of Biondi (2001). In particular, ray theory breaks down if its high frequency assumption does not hold (Woodward, 1990), as is the case with velocity anomalies of the size of those that generate the FEAVO effect. Multipathing is another instance when the wave equation methods are more appropriate. They also treat more correctly the amplitudes.

WEMVA proceeds as follows: the wavefield at a certain depth is downward continued a depth step through the known velocity model with an accurate (nonlinear) operator. The result is transformed to angle domain and an improved image is created by eliminating either the curvature of the events in angle-domain common image gathers[2] (Prucha et al., 1999), or the FEAVO anomalies. An image perturbation (ΔImage) is obtained by subtracting the two images, and is backprojected through an invertible operator in order to obtain a velocity update. The velocity model is updated and the cycle proceeds again until ΔImage becomes negligible. Figure 1 illustrates this process.

WEMVA is at the forefront of research. As the comparison in the next section will show, it avoids all the shortcomings of the previous attempts.

THE NEW APPROACH

My goal is to modify a brand-new tool (WEMVA) to the specific of an old unsolved problem (velocity model finding for FEAVO elimination). The FEAVO anomalies will be eliminated by finding an accurate velocity model, then by downward continuing the data through it After obtaining a velocity model good enough to help eliminate the FEAVO, the significance of the process will be tested by performing geological interpretations of the data with and without the FEAVO anomalies.

The strategy of WEMVA differs from that of all the previous attempts: instead of trying to fit the data, it tries to fit the image. The tactics are different as well: instead of integral (ray-based, Kirchhoff) operators, it uses differential (wave-equation) ones, with all their well-

[2]referred from now on simply as "angle gathers," or by the acronym ADCIGs

known advantages. Unlike most of the previous approaches, this one is capable of using a starting guess from classic velocity analysis. Therefore, it simply fine tunes the velocity model for small velocity anomalies. It requires no picking and tries to match the entire image using information contained both in the amplitudes and traveltimes.

This new approach, unlike the old ones, takes into account all the characteristics of the FEAVO anomalies, including the variation with depth. This goal is achieved by measuring anomalies on surfaces in the depth-midpoint-angle volume, instead of simple "V"s in the midpoint-offset plane. An adaptation of the signal-noise separation technique described in Harlan (1986) will assure that the image perturbation contains only information related to the FEAVO anomalies.

My approach uses the strategy and framework of WEMVA as described in the corresponding section, but differs by the change of objective function in the inversion. The usual WEMVA criterion describing the quality of the image is flatness in angle gathers that is directly related to traveltime anomalies. The traveltime changes associated with the FEAVO effect are very small and do not produce whole-event curvature, but only wiggliness 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 the characteristic "V" patterns in the midpoint-offset plane.

My method might be able to discriminate between amplitude anomalies caused by absorption and those caused by velocity because both kinds of anomalies have two different "signatures" in the image space. In the case of velocity, the high amplitudes are found close to the low amplitudes: the energy is not lost but it is only focused locally. In the case of absorption the FEAVO effects are not "bipolar."

EVIDENCE THAT THE PROPOSED APPROACH IS FEASIBLE

As stated by Vlad and Biondi (2002), the heuristic for proving that the inversion approach is feasible comprises three steps:

1. Proving that shallow-origin FEAVO anomalies are visible in the angle domain and wave-equation methods correctly handle the deep-origin FEAVO effects in synthetic data. That was first accomplished by Vlad and Biondi (2002) and in this paper I present some improved results.

2. Proving that the anomalies can be extracted in the angle domain. I will show that in the case of simple velocity distributions, the shape of the anomalies is described by very simple analytical formulas and I will discuss ways to extract them.

3. Proving that the extracted anomalies can be transformed back to the initial domain, updating the guess. In particular, that the linearized downward continuation does not render FEAVO effects unrecognizable.

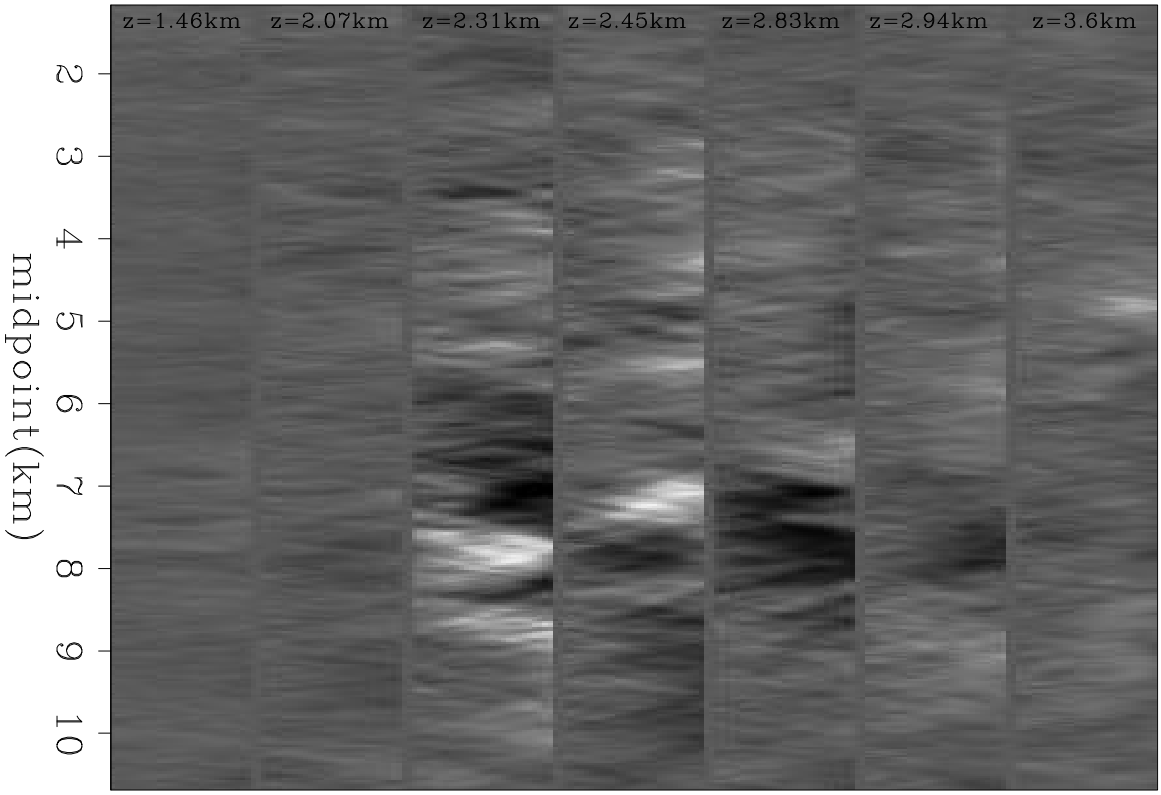


Figure 2: Midpoint-angle slices, reprocessed dataset, new velocity analysis. The improvement in quality allowed the panels to go up to 25 degrees, instead of only 20 in the previous paper. nick2-fkapanel [CR]

Shallow-origin FEAVO effects

Vlad and Biondi (2002) have proven this point using the very dataset on which FEAVO was first observed by Kjartansson (1979); however, its preprocessing and preliminary velocity analysis left a lot of room for improvement. The details of the preprocessing are described and illustrated in Appendix C. A new velocity analysis, also illustrated in Appendix C, follows. The result of phase-shift migration and conversion to angle gathers is visible in Figure 2.

Deep-origin FEAVO effects

Hatchell (2000) proves using both real and synthetic data that FEAVO effects can be generated not only by shallow velocity anomalies, but also by deep ones. They usually have a different origin than the shallow ones. They appear not because of depositional irregularities, but because of sudden terminations of thin layers with anomalous velocity against subvertical faults. The question of whether the deep-origin FEAVO behaves like its shallow counterpart and are visible in angle gathers is, therefore, legitimate.

The easiest way to answer this would be to repeat the same experiment as in the previous

subsection, but real data containing deep-origin FEAVO is not yet available. Therefore, I will use synthetic data to show that wave-equation methods handle well the deep-origin FEAVO effects. These results are in Appendix A.

Extracting FEAVO effects in angle gathers

Extracting the FEAVO anomalies assumes that the long spatial wavelength velocity model is good enough that the reflectors are flat in angle gathers, and only the amplitudes remain to be fixed. Since the FEAVO effects are expressed both in the midpoint-angle domain and in the angle-domain common image gathers, their separation must proceed in a synergistic fashion. This involves the entire data volume. For the simplest case (constant velocity, flat reflectors), the FEAVO effects generated by a velocity anomaly at depth z_a and midpoint m_a will be distributed in the depth(z)–midpoint(m)–angle(θ) space along a surface described by:

$$z = z_a + |m - m_a| \cot\theta. \tag{1}$$

The derivation is laid out in Appendix B, and the shape of the surface is shown in Figure 3. Even for a v(z) case (Grand Isle dataset), the shape of the anomaly will not be very different,

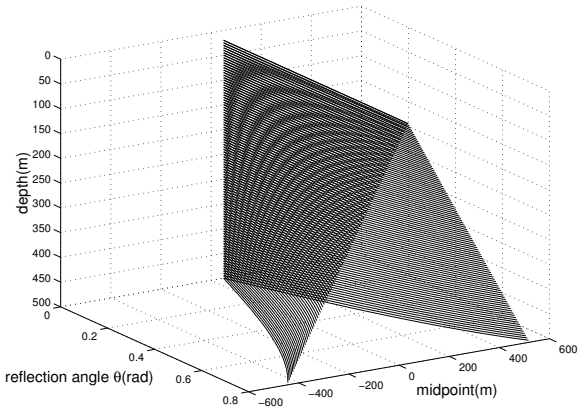


Figure 3: Shape of the FEAVO "footprint" in the depth-midpoint-angle space due to a velocity anomaly 20 m deep in an otherwise constant velocity medium with flat reflectors. For a better 3D visual understanding, the shape resembles the bow of a flipped boat. nick2-ang3d_20_500_pi4 [NR]

especially for a limited angle range (due to a finite range of offsets). Figure 4 shows the dips of the FEAVO effects are confined to a limited range. Therefore, as a first measure for separating them, we can apply an appropriate f-k dip filter to the midpoint–angle slices (Figure 2). This eliminates the largest part of the petrophysical AVO. There is, however, no guarantee that the remaining energy within the plausible FEAVO dip range does actually belong to FEAVO. I will have to separate the signal from noise in the manner of Harlan (1986):

1. For each point in the depth-midpoint section, consider that it "houses" an anomaly and precompute the FEAVO-effect surface that depends on the known long spatial wavelength velocity field.

2. Sum (or compute a semblance-like operator) along the precomputed surfaces to obtain a depth-midpoint "anomaly map," taking care to distinguish between FEAVO caused by absorption and that caused by velocity.

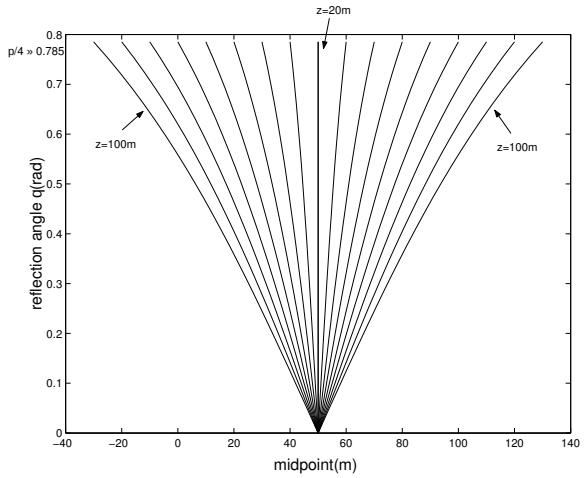


Figure 4: Midpoint–angle contour map of FEAVO effects generated by a velocity anomaly 20m deep. The angle range is wide (up to 45°), which is wider than the range recorded in most of the real data sets. Therefore, it is unlikely that curvature of the anomalies be observed in real data panels. nick2-ang20_100_pi4 [NR]

3. Filter the image based on its statistical properties, so only the most focused points remain.

4. Spread the filtered image along the precomputed surfaces back into the depth-midpoint-angle space. Alternately, focusing could be done using the downward continuation operator itself.

Linearized downward continuation preserves the FEAVO effects

The third WEMVA step relies on inverting a linearized downward continuation operator in order to obtain the velocity perturbation from the image perturbation. This means the linearization must not destroy the FEAVO anomalies. One way to check this is to actually do WEMVA for a synthetic case. A smarter, less time-consuming way is to see whether the non-inverted operator correctly propagates a wavefield through a velocity anomaly to create a FEAVO effect.

A good comparison case can be provided by the waveform modeling of deep FEAVO anomalies (Figure 6). Figure 5 represents the results of an equivalent experiment - propagating a shot (20Hz Ricker wavelet, laterally smoothed a bit) from the surface to a line of receivers 6 km deep. The difference is that in Figure 5 the propagation was done with linearized downward continuation [the complexified local Born-Fourier method (de Hoop et al., 2000), as described by Sava (2000)], instead of pseudospectral waveform modeling. Details about the operator and the way the image was constructed are in Appendix B. The FEAVO effects are easily recognizable in amplitudes and the dispersion is missing. Even if they are less powerful than in Figure 6, especially in time, they are clearly distinguishable.

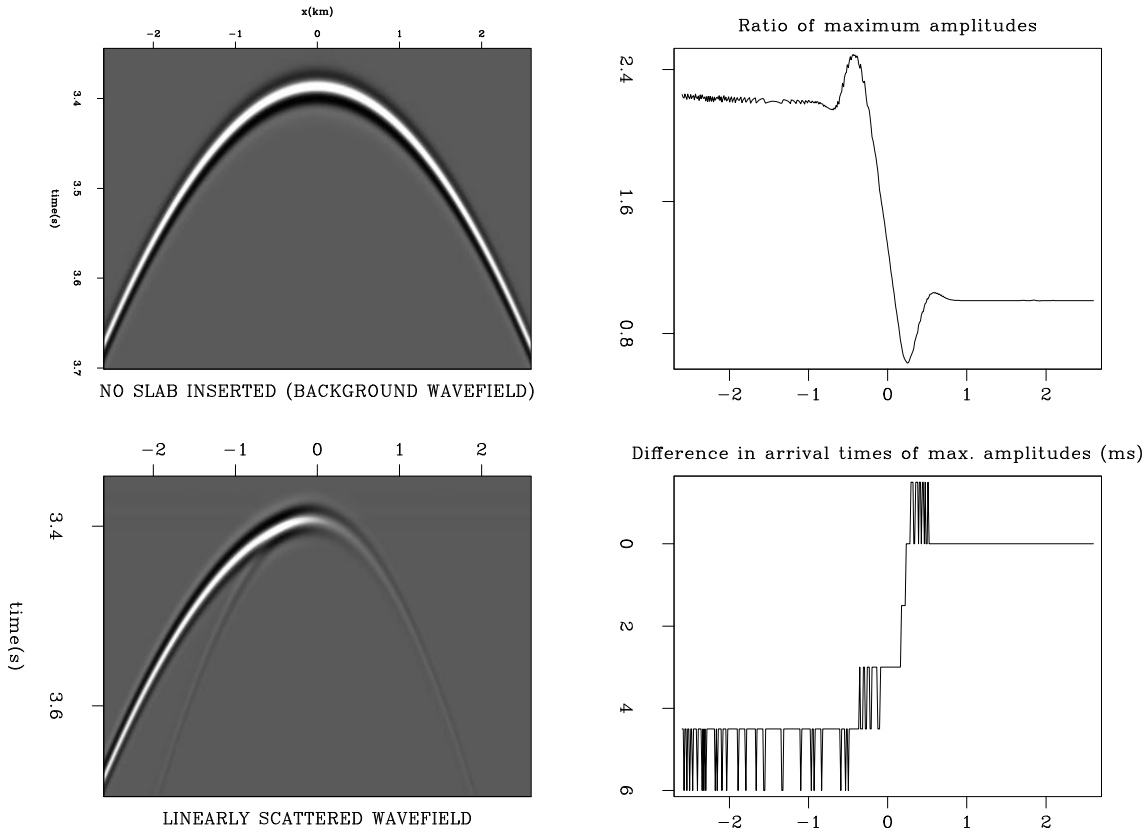


Figure 5: **Left, from top to bottom:** 1. Wavefield recorded 6 km deep after propagation through constant velocity (background wavefield); 2. Linearly scattered wavefield (physically equivalent to the difference between the wavefield propagated through the velocity model containing the slab – panel 6 of Figure 6 – and the background wavefield); **Right, from top to bottom:** 3. Ratio between the maximum amplitudes in panel (1+2) and panel 1, for each x location; 4. Difference between the times of the maximum amplitudes in 1 and (1+2), for each x location. The wavefield was propagated by linearized downward continuation (complexified local Born-Fourier method) instead of pseudospectral waveform modeling. nick2-popic [ER]

CONCLUSIONS

Building a velocity model accurate enough to remove FEAVO effects by prestack depth migration is an important problem which has not been solved satisfactorily until now. WEMVA, with a fitting goal adapted to the specific morphology of the anomalies, is a promising tool that will likely solve the problem. Work in this direction has proven that the iterative inversion is feasible: real data focusing effects are visible in the angle domain, the anomalies are preserved by waveform modeling, and the linearized downward continuation operator does not destroy the anomalies.

FUTURE WORK

The previous sections have proven that an inversion for the velocity model that produces the FEAVO anomalies is possible. The actual work of setting up such an inversion first for 2D, then for 3D, remains to be done. While the functioning of the forward and the inverse operators for the iterative inversion is being proved by the work of Paul Sava, I have to replace his approach of constructing $\Delta Image$ with my approach of extracting the FEAVO anomalies.

I will also need to investigate ways to discriminate between absorption and velocity caused FEAVO. I plan to study the effects of the source directivity on the amplitudes and to investigate a non-smoothing styling goal for the inversion. I would have to investigate ways to do surface-consistent amplitude corrections that will account for surface absorption variations, and to see whether that will not destroy the FEAVO.

Finally, I will perform geological interpretations of the data with FEAVO removed, and compare them with interpretations of the original data. This should show that FEAVO removal and the new velocity model made a difference in interpretation results.

ACKNOWLEDGMENTS

I am grateful to Biondo Biondi for providing me strategic guidance pertaining to the approach to the problem and for the waveform modeling in Figure 6, to Paul Sava and Bob Clapp for their constant good advice and support, and to Rick Ottolini and Andrey Karpushin for bringing relevant references to my attention.

REFERENCES

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

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.

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

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

Clapp, R. G., Sava, P., and Claerbout, J. F., 1998, Interval velocity estimation with a null-space: SEP–**97**, 147–156.

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

de Hoop, M. V., le Rousseau, J. H., and Wu, R.-S., 2000, Generalization of the phase-screen approximation for the scattering of acoustic waves: Wave Motion, **31**, no. 1, 43–70.

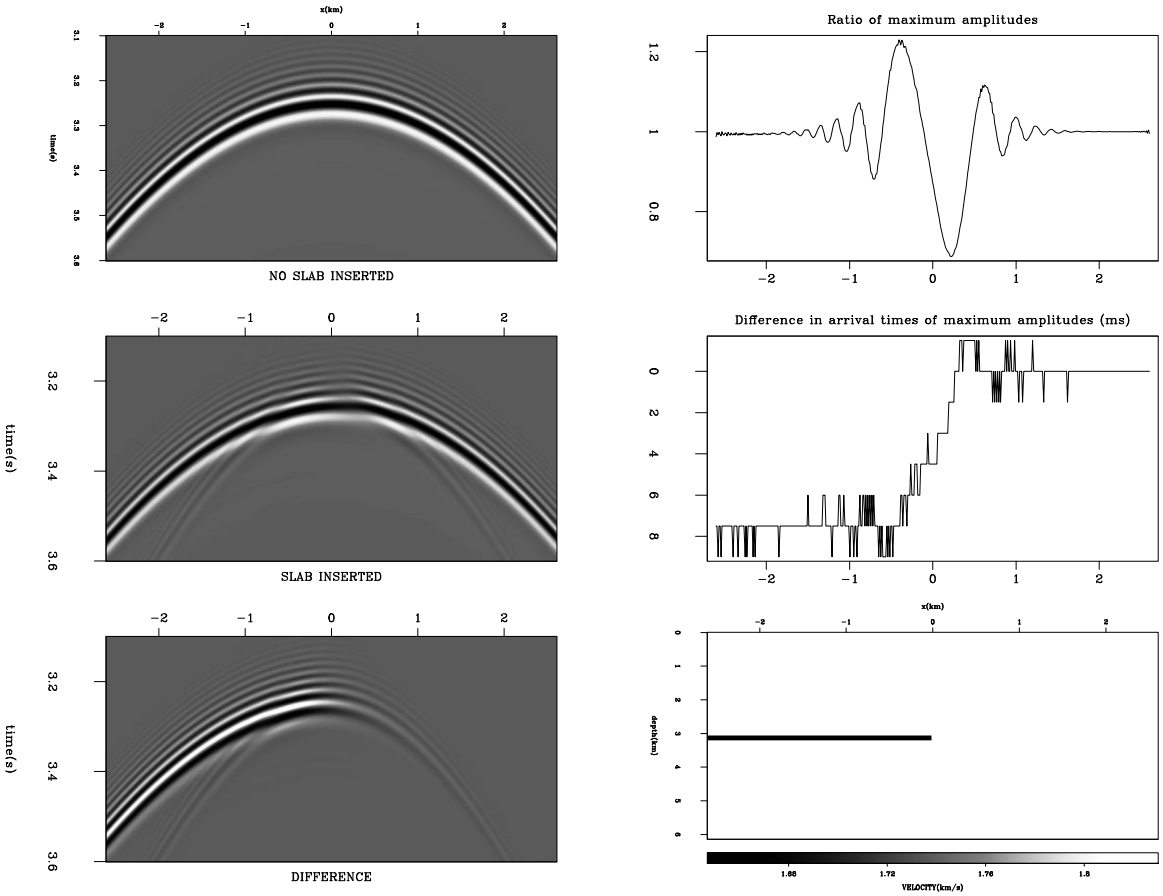


Figure 6: **Left, from top to bottom:** 1. Wavefield recorded 6 km deep after propagation through constant velocity; 2. Wavefield recorded 6 km deep after propagation through velocity model in panel 6; 3. Difference between 1 and 2; **Right, from top to bottom:** 4. Ratio between the maximum amplitudes in panel 2 and panel 1, for each x location; 5. Difference between the times of the maximum amplitudes in 1 and 2, for each x location; 6. Velocity model for panel 2 – homogeneous with a lower velocity slab inserted. nick2-hatsim [CR]

Harlan, W. S., 1986, Signal-noise separation and seismic inversion: Ph.D. thesis, Stanford University.

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

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

Ottolini, R., and Rocca, F., 1982, Direct observation of lateral velocity anomalies in field data: SEP–**32**, 15–23.

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

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

Vlad, I., and Biondi, B., 2002, Velocity estimation for seismic data exhibiting focusing-effect avo: SEP–**111**, 107–123.

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

Woodward, M., 1987, Reflection tomography : Vees in midpoint offset space: SEP–**51**, 1–12.

Woodward, M., 1990, Wave equation tomography: Ph.D. thesis, Stanford University.

APPENDIX A – DERIVATION OF FORMULAS FOR FEAVO ANOMALY SHAPES

All the derivations that follow are made under the assumption that the velocity is constant (straight rays) and that all reflectors are flat and horizontal. In order to derive the shape of FEAVO effects in the angle domain (1), we first derive the shape of FEAVO effects in the offset domain. Consider case B (the general case) in Figure 7 Vlad and Biondi (2002). For the zero-offset experiment, the focusing-generating anomaly affects only its own midpoint. For any other offsets, it affects two midpoints that grow increasingly distant with offset. In Figure 8, because the reflector is parallel to the surface,

$$\left.\begin{array}{c} CD||AE \Rightarrow B\hat{D}C = D\hat{A}E \\ DCB = AED = 90^\circ \end{array}\right\} \Rightarrow \Delta BDC \approx \Delta DAE \Rightarrow \frac{BC}{DE} = \frac{DC}{AE} \Rightarrow \frac{z - z_a}{z} = \frac{m - m_a}{f/2} \tag{2}$$

Applying the same reasoning to the left side of case B in Figure 7, we can write the equation for both slanted streaks at depth z as

$$f = \frac{z}{z - z_a} \cdot 2\,|m - m_a| \tag{3}$$

Figure 9 depicts parts of the corresponding surface for a 20m deep anomaly. Notice the arched form of the surface with the midpoint–depth vertical planes at maximum offset. This (with very different vertical scaling) is the "bullet shape" observed by Ottolini and Rocca (1982) in a real dataset. The offset f can be easily replaced with the reflection angle in this case because the reflector is flat:

$$\theta = A\hat{D}E \Rightarrow \tan\theta = \frac{AE}{ED} = \frac{f}{2z}. \tag{4}$$

Plugging in (3),

$$|m - m_a| = (z - z_a)\tan\theta \tag{5}$$

which can also be written as (1).

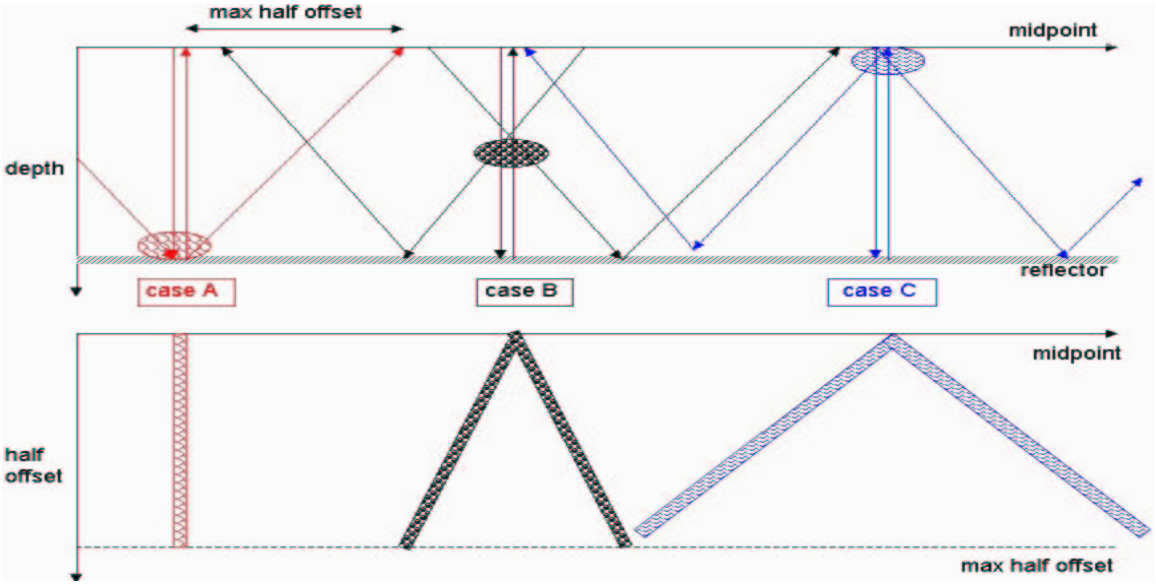


Figure 7: Physical explanation for the expression of FEAVO anomalies in midpoint-offset space. In the upper picture, the blobs are transmission anomalies and the arrows are raypaths for zero offset and maximum offset recordings. For case A (anomaly on the reflector), only a single midpoint is affected 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]

APPENDIX B – THE FORWARD LINEARIZED DOWNWARD CONTINUATION OPERATOR

Downward continuation can be done in the Fourier domain as a phase shift applied to U_z, the wavefield at depth z (Claerbout, 1985):

$$U_{z+\Delta z} = e^{ik_z \Delta z} U_z \tag{6}$$

The vertical wavenumber k_z depends on the laterally varying velocity. In order for us to be able to implement the formula, we have to decompose k_z into a part not influenced by lateral velocity variations (k_{z_o}) and a part influenced by them (k_{z_x}):

$$k_z = k_{z_o} + k_{z_x} \Rightarrow U_{z+\Delta z} = e^{i(k_{z_o}+k_{z_x})\Delta z} U_z \Rightarrow U_{z+\Delta z} = e^{ik_{z_o}\Delta z} e^{ik_{z_x}\Delta z} U_z \tag{7}$$

The Born approximation is equivalent to a linearization of the exponential $e^x \approx 1 + x$, and therefore

$$U_{z+\Delta z} \approx e^{ik_{z_o}\Delta z} U_z (1 + ik_{z_x}\Delta z) \tag{8}$$

In the case of the complexified local Born-Fourier (complexified pseudo-screen) method, with the notations in equations (12) and (13) of Sava (2000), we can rewrite it as:

$$U_{z+\Delta z} \approx \mathcal{T} U_z [1 + \mathcal{S}(s - s_o)] \tag{9}$$

where $\mathcal{T}$ is the background wavefield downward continuation operator applied in the $\omega - \mathbf{k_m}$ domain:

$$\mathcal{T} = e^{i\Delta z\sqrt{\omega^2 {s_o}^2 - (1-i\eta)^2|\mathbf{k_m}|^2}}, \tag{10}$$

$\mathcal{S}$ is the scattering operator, applied in the $\omega - x$ domain:

$$\mathcal{S} = \frac{i\Delta z\omega^2 s_o}{\sqrt{\omega^2 {s_o}^2 - (1-i\eta)^2|\mathbf{k_m}|^2}} \tag{11}$$

and where s is the slowness at the depth $z+\Delta z$, $\mathbf{k_m}$ is the wavenumber across the midpoint direction (scalar for 2D, vector for 3D), s_o is the constant background slowness, ω is the frequency, and η is a small dimensionless quantity introduced for numerical stability; $\mathbf{k_m}$ and ω must contain a 2π constant. The output of this operator can be seen in Figure 10.

Although with equation (9) we went a step towards linearity with respect to the slowness perturbation term, it is not fully linear because the slowness perturbations compose with themselves. This is visible if we examine the first two steps of the downward continuation. At $z=0$, $U_{z=0} = Data$ (Ricker wavelet at zero-time in the middle of the x axis). At $z=\Delta z$,

$$U_{z=\Delta z} = \mathcal{T}U_{z=0} + \mathcal{S}\Delta s_{z=\Delta z}\mathcal{T}U_{z=0}. \tag{12}$$

At $z=2\Delta z$,

$$U_{z=2\Delta z} = \mathcal{T}U_{z=\Delta z} + \mathcal{S}\Delta s_{z=\Delta z}\mathcal{T}U_{z=\Delta z}, \tag{13}$$

and by plugging in the expression for $U_{z=\Delta z}$ and because $\mathcal{T}$ and $\mathcal{S}$ do not commute,

$$U_{z=2\Delta z} = \mathcal{T}\mathcal{T}U_{z=0} + \mathcal{T}\mathcal{S}\Delta s_{z=\Delta z}\mathcal{T}U_{z=0} + \mathcal{S}\Delta s_{z=2\Delta z}\mathcal{T}\mathcal{T}U_{z=0} + \mathcal{S}\Delta s_{z=\Delta z}\mathcal{T}\mathcal{S}\Delta s_{z=2\Delta z}\mathcal{T}U_{z=0} \tag{14}$$

In order to obtain a downward continuation that is linear in the slowness perturbations Δs, we have to drop the last term at each step. Thus, after the n^{th} depth step, the wavefield will be:

$$U_{z=n\Delta z} = \left(\prod_1^n \mathcal{T}\right)U_{z=0} + \sum_{j=1}^{n}\left[\left(\prod_1^{n-j}\mathcal{T}\right)\mathcal{S}\Delta s_{z=j\Delta z}\left(\prod_1^{j}\mathcal{T}\right)U_{z=0}\right] \tag{15}$$

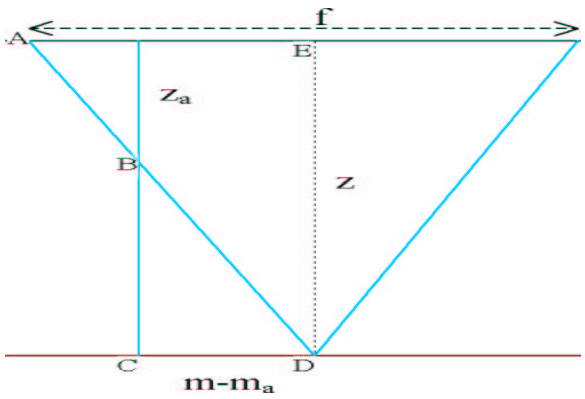


Figure 8: The right half of case B in Figure 7. Raypaths are in blue. The transmission anomaly is in B, at a depth of z_a. AE (of length f – full offset) is at the Earth's surface, C is on the reflector at the anomaly midpoint (m_a), D is on the reflector at midpoint m and depth z. DE is perpendicular to the surface; BC is perpendicular to the reflector. nick2-skema2 [NR]

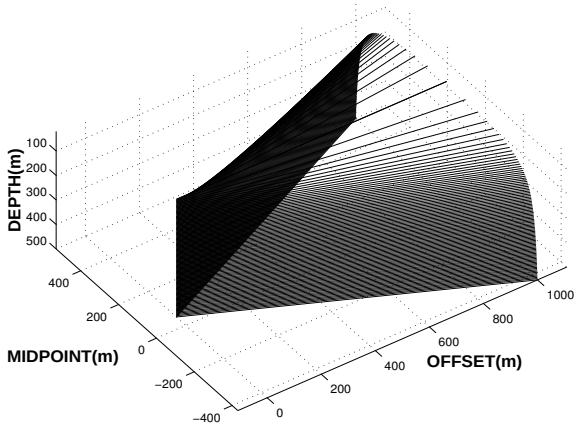


Figure 9: Fragment of the surface described by equation 3, between 0 and 500m, for a transmission anomaly 20 m deep. The shape resembles the bow of an overturned boat. nick2-20_max500 [NR]

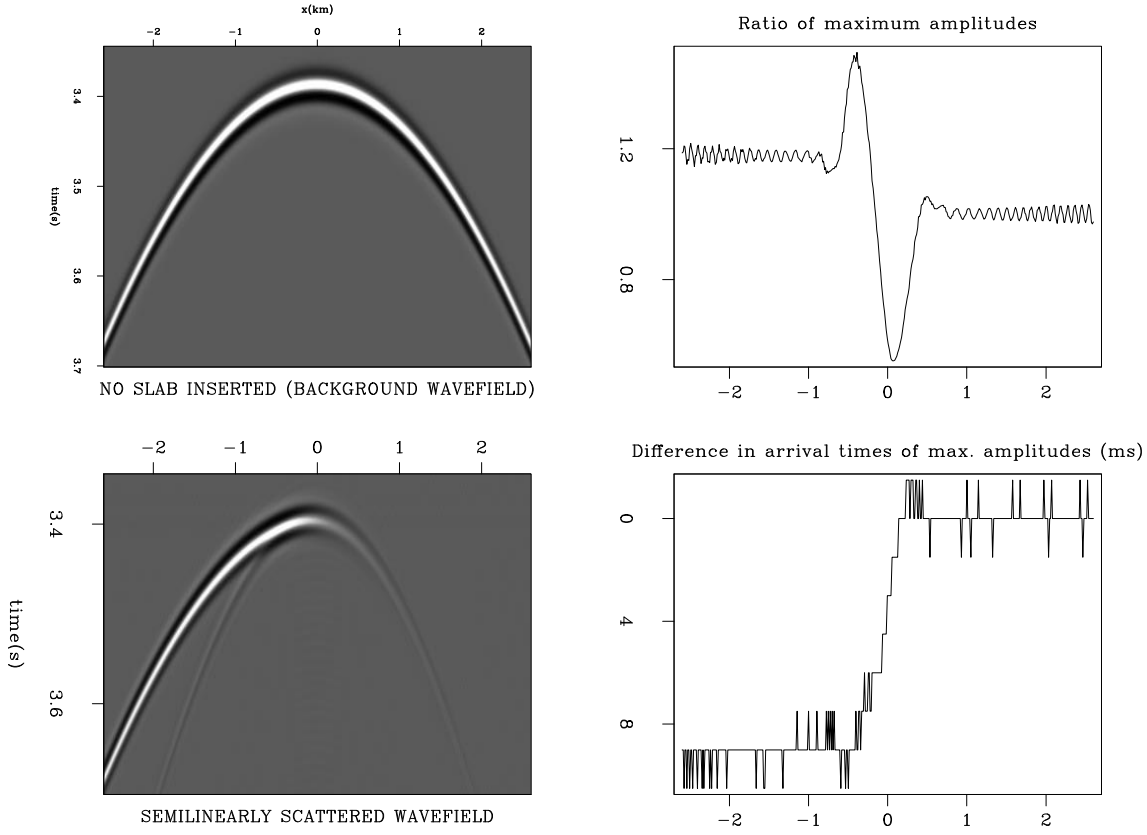


Figure 10: **Left, from top to bottom:** 1. Wavefield recorded 6 km deep after propagation through constant velocity (background wavefield); 2. Quasi linearly scattered wavefield (physically equivalent to the difference between the wavefield propagated through the velocity model containing the slab – panel 6 of Figure 6 – and the background wavefield); **Right, from top to bottom:** 3. Ratio between the maximum amplitudes in panel (1+2) and panel 1, for each x location; 4. Difference between the times of the maximum amplitudes in 1 and (1+2), for each x location. The wavefield was propagated with the operator in equation 9. Panel 4 is in very good accord with panel 4 of Figure 6 and with the analytical time delay (8.7 ms). nick2-patpic [ER]

The above formula is equivalent with stating that at each level, we compute the scattered wavefield only from the background wavefield from the previous level, then we propagate it down until the last level with the background operator. The results of this approach are visible in Figure 5.

APPENDIX C - PREPROCESSING DETAILS

I first applied better tuned f-k filters, then shifted the data 9 meters across offset using a frequency-domain operator. Why? The migration program `Phase` requires data to be regularly sampled to contain the zero offset. The minimum offset of the data was 241m and the offset sampling was 50m (interpolated to 25), so there was no way of having both the zero offset and regularly sampled data. Worse, `Phase` requires split-spread data, so half of the offsets would have been off by 9 meters. I then performed f-x decon to eliminate random noise. I interpolated the offsets from a sampling rate of 50m (visible aliasing) to 25m in the wavenumber domain. I performed deconvolution using `Pef` and `Helicon`. I had to apply again f-k filters with new parameters to eliminate some of the effects of former aliasing, which turned into spurious events after interpolation.

Figures 11 and 12 show the smallest non-extrapolated offset before and after the new preprocessing, respectively. The railroad-track reflections above 1.5 seconds, which is actually water-velocity noise, is eliminated and the geology beneath is uncovered (due to the dip filters). The strong ringing which multiplied reflectors most visibly in the high-amplitude region is gone (due to deconvolution). The signal/noise ratio between 3 and 5 seconds is highly improved (due to the f-x decon). After the new preprocessing, the stratigraphy looks much more interpretable and new, subtler FEAVO anomalies are brought to light. The V-shaped anomalies were not destroyed; on the contrary, they are clearer than ever (Figure 13).

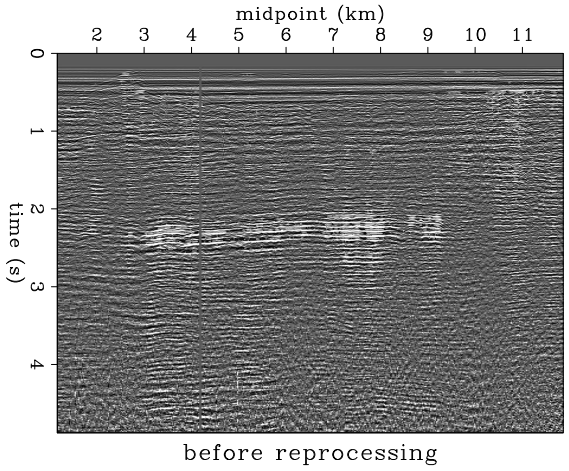


Figure 11: Smallest offset (241m) before reprocessing nick2-zofbef [CR]

The previous velocity model, which is already existing in the data library, is shown in the upper left panel of Figure 14. The geological setting of the Grand Isle survey in the Mississippi Delta shows that the Grand Isle deposits are very young and the velocity is most likely

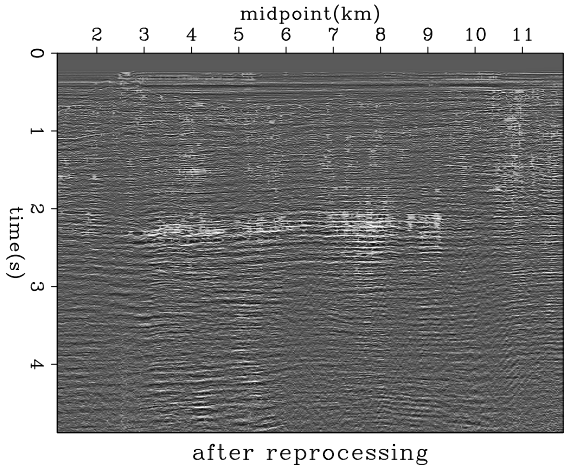


Figure 12: Smallest offset (250m) after reprocessing. Railroad-track false reflections above 1.5 sec, ringing all over the section and high noise in the lower part are eliminated. nick2-zofaft [CR]

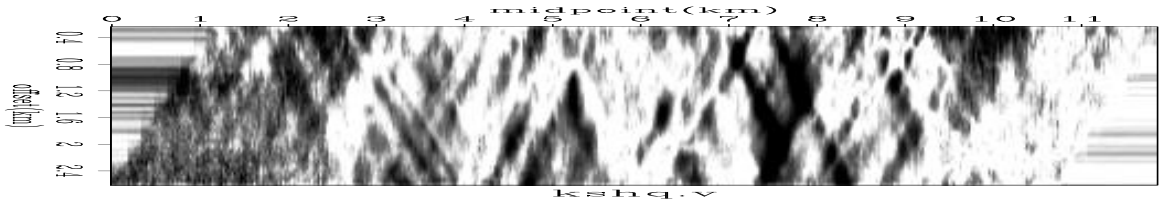


Figure 13: Preprocessing enhanced the V-shaped anomalies nick2-kshq [CR]

determined by compaction, making such large lateral velocity variations as pictured in the initial model implausible. The previous velocity had also been picked at only ten midpoints.

I eliminated random noise from the data with an enhanced noise attenuation method. I then transformed each CMP to velocity space, automatically picked the highest semblance values, and transformed them to interval velocity using the "SuperDix" inversion described by Clapp et al. (1998) (Figure 15). The result of the inversion was then smoothed along midpoint into a more geologically plausible almost-v(z).

I migrated with the velocity shown in the lower left panel of 14. I also used more frequencies than in the previous migration. The new migration stack is shown in Figure 16. Some reflectors stack better in the newer result, and amplitude anomalies are also more consistent.

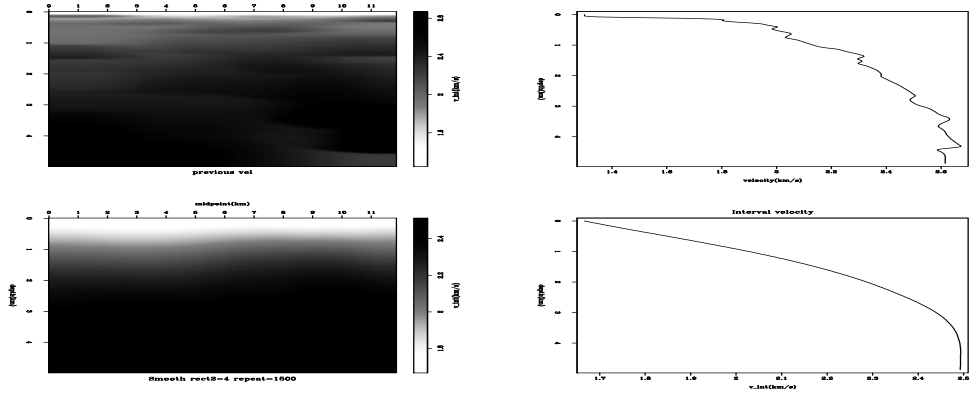


Figure 14: Upper left: previous interval velocity model. Upper right: v(z) model constructed by smoothing it many times. Lower left: new interval velocity model for migration. Lower right: "v(z)" profile constructed by smoothing the new velocity model across midpoint nick2-veloplot [CR]

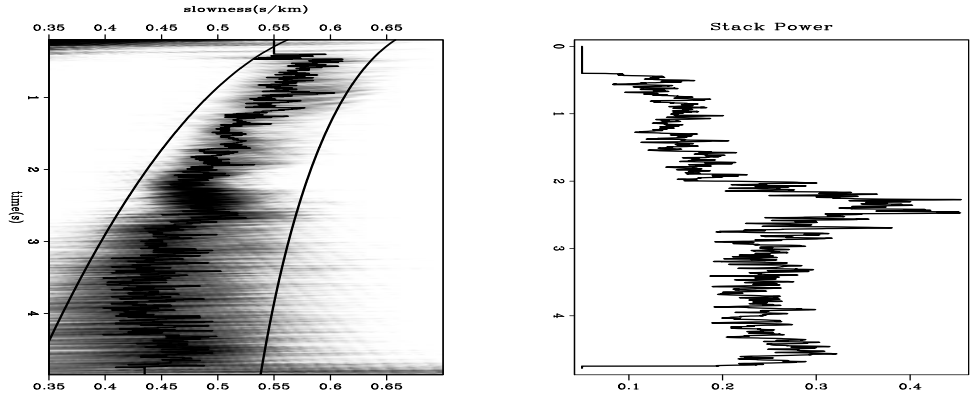


Figure 15: Illustration of the velocity analysis for one midpoint: autopicker fairway, automatic picks, and inversion weights. nick2-phw [CR]

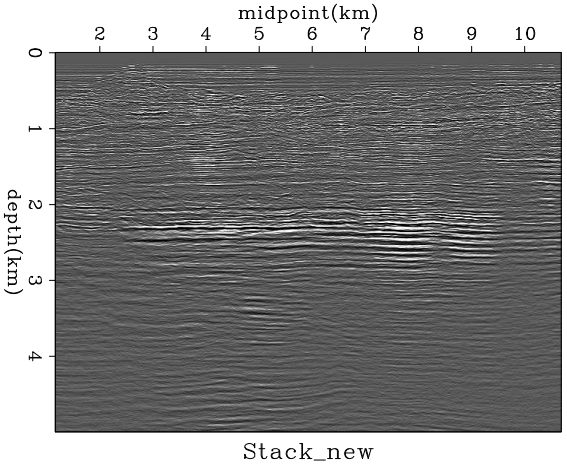
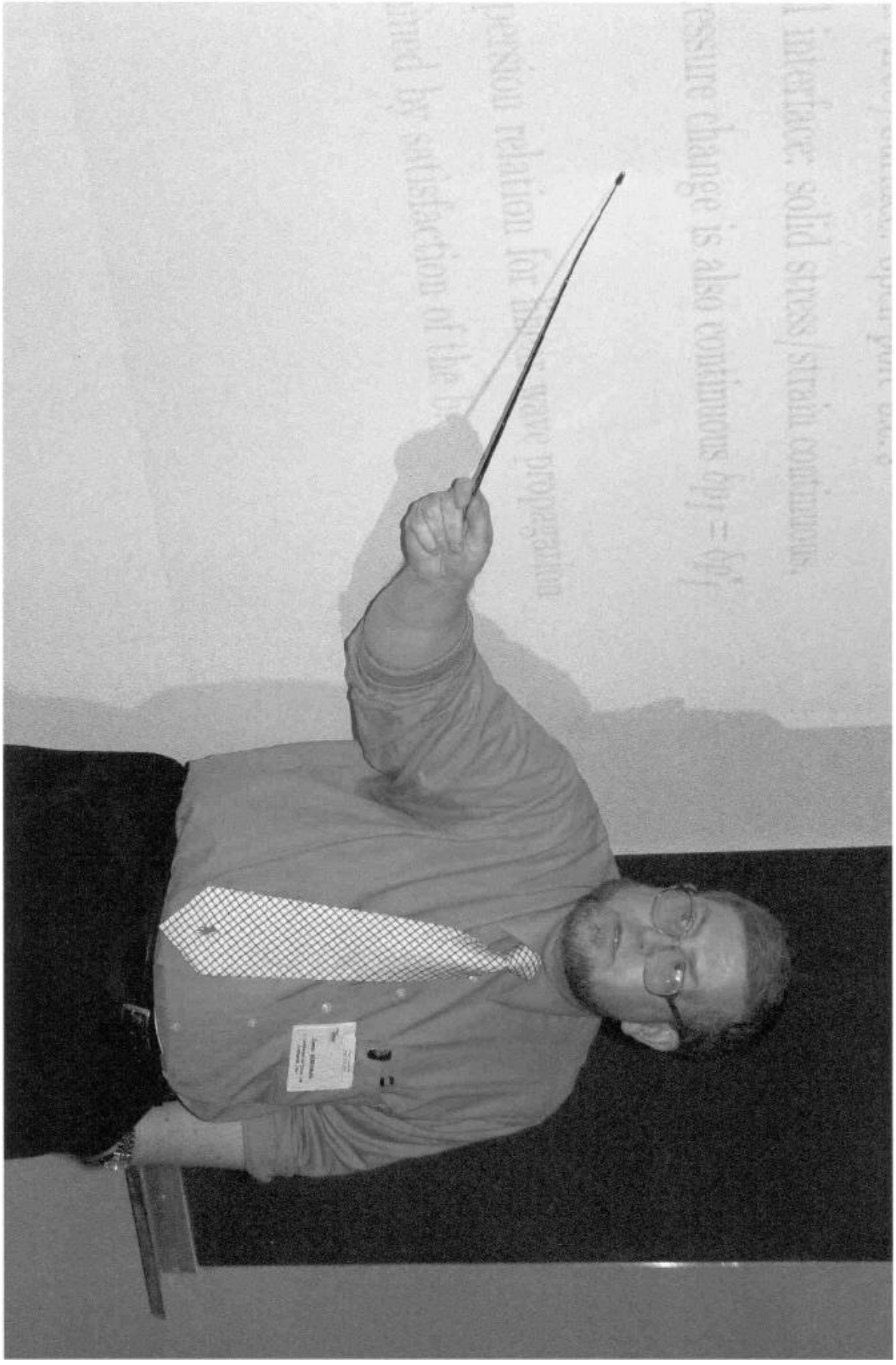


Figure 16: New migrated stack nick2-kaer_new [CR]



Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 65–83

Transformation to dip-dependent Common Image Gathers

Biondo Biondi and William Symes[1]

ABSTRACT

We introduce a new transform of offset-domain Common Image Gathers (CIGs) obtained by wavefield-continuation migration methods. This transformation can be applied to either horizontal-offset CIGs or vertical-offset CIGs. It overcomes the limitations that both kinds of CIGs suffer in the presence of a wide range of reflectors' dips. The result of our transformation is an image cube that is equivalent to the image cube that would have been computed if the offset direction were aligned along the apparent geological dip of each event. The proposed transformation applies a non-uniform dip-dependent stretching of the offset axis and can be efficiently performed in the Fourier domain. Because it is dependent on the image's apparent dip, the offset stretching automatically corrects for the image-point dispersal. Tests on a synthetic data set confirm the potential advantages of the transformation for migration velocity analysis of data containing steeply dipping reflectors.

INTRODUCTION

The analysis of Common Image Gathers (CIG) is an essential tool for updating the velocity model after depth migration. When using wavefield-continuation migration methods, angle-domain CIGs (ADCIGs) are usually used for velocity analysis (Clapp and Biondi, 2000). The computation of ADCIG is based on slant-stack transformation of the wavefield either before imaging (Prucha et al., 1999) or after imaging (Sava et al., 2001; Rickett and Sava, 2001; Biondi and Shan, 2002). In either case, the slant stack transformation is usually applied along the horizontal offset axis.

However, when the geological dips are steep, this "conventional" way of computing CIGs does not produce useful gathers, even if it is kinematically valid for all geological dips that are milder than 90 degrees. As the geological dips increase, the horizontal-offset CIGs (HOCIGs) degenerate, and their focusing around zero offset blurs. This limitation of HOCIGs led both of the authors to independently propose a partial solution to the problem; that is, the computation of CIG along a different offset direction than the horizontal one, and in particular along the vertical direction (Biondi and Shan, 2002). Unfortunately, neither set of angle-domain gathers (HOCIG ad VOCIG) provides useful information for the whole range of geological dips, making their use for velocity updating awkward. While VOCIG are a step in the right direction, they are not readily usable for migration velocity analysis (MVA).

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

In this paper we present a new method to transform a set of CIGs (HOCIGs and/or VOCIGs) into another set of CIGs. The resulting image cube is equivalent to the image cube that would have been computed by aligning the offset direction along the local geological dips. This transformation applies a non-uniform dip-dependent stretching of the offset axis and can be cheaply performed in the Fourier domain. Because the offset stretching is dependent on the reflector's dip, it also automatically corrects for the image-point dispersal. It thus has the potential to improve substantially the accuracy and resolution of residual moveout analysis of events from dipping reflectors. It has been recognized for long time (Etgen, 1990) that image-point dispersal is a substantial hurdle in using dipping reflections for velocity updating.

The proposed transformation is dependent on the apparent dips in the image cube, and creates an image cube in which the effective offset depends on those apparent dips in an "optimal" way. We will thus refer to the resulting CIGs as dip-dependent offset CIGs (DDOCIGs), and to the transformation as the "transformation to DDOCIGs."

The proposed method is independent from the particular migration method used to obtain the CIGs. The input offset-domain image cube can be computed by either downward-continuation migration (shot profile or survey sinking) or reverse-time migration (Biondi, 2002). The proposed transformation should also improve the accuracy and resolution of velocity analysis applied only to "conventional" HOCIGs. In its most immediate application, it should also improve the image obtained by stacking after a residual moveout correction.

The next section illustrates the problem of HOCIGs and VOCIGs with a real data set from the North Sea that was recorded above a steeply dipping salt edge. The following section introduces the new transformation, that is then tested on a synthetic data set.

COMMON IMAGE GATHERS AND STRUCTURE: A NORTH SEA EXAMPLE

Figures 1–3 illustrate the problem with HOCIGs for a North Sea data set where the salt body has a vertical edge. Because of presence of overturned paths, the data were imaged using a shot-profile reverse time migration.

Figure 1 shows the image extracted at zero offset, which is equivalent to the "stacked image" for Kirchhoff-like migration methods. The vertical edge is well imaged at zero offset, but when we analyze the image as a function of offset at the fixed surface location corresponding to the vertical salt edge (right panel in Figure 2), we immediately notice that, at the depth interval corresponding to the salt edge, there is no focusing along the offset axis. In contrast, the focusing along offset is obvious when we analyze the image at the surface location corresponding to mild reflector dips (left panel in Figure 2). As expected, the lack of focusing in the HOCIGs carries over to the image after transformation to angle domain by slant stacking (Figure 3). In the next section we will explain the degradation of the horizontal-offset CIGs by a simple geometric analysis.

Figure 5 shows a vertical-offset CIG (VOCIG) for the same data set. Since the offset is vertical, the image cube is cut at a constant depth, not at constant surface location. The depth of this CIG corresponds to the black line superimposed onto the image in Figure 4. Now the

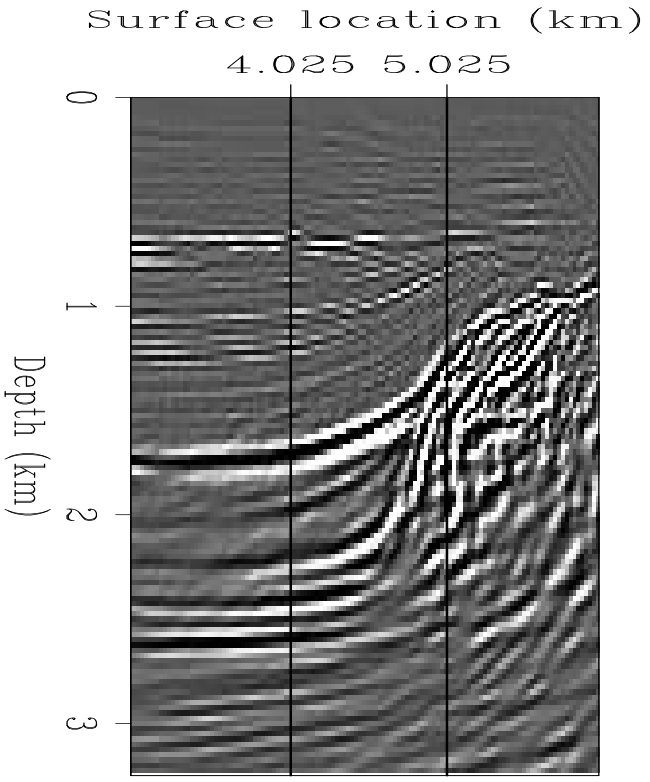


Figure 1: Image of the North Sea data set. The black lines superimposed onto the image indicate the positions of the HOCIGs shown in Figure 2. biondo1-image-cig-new [NR]

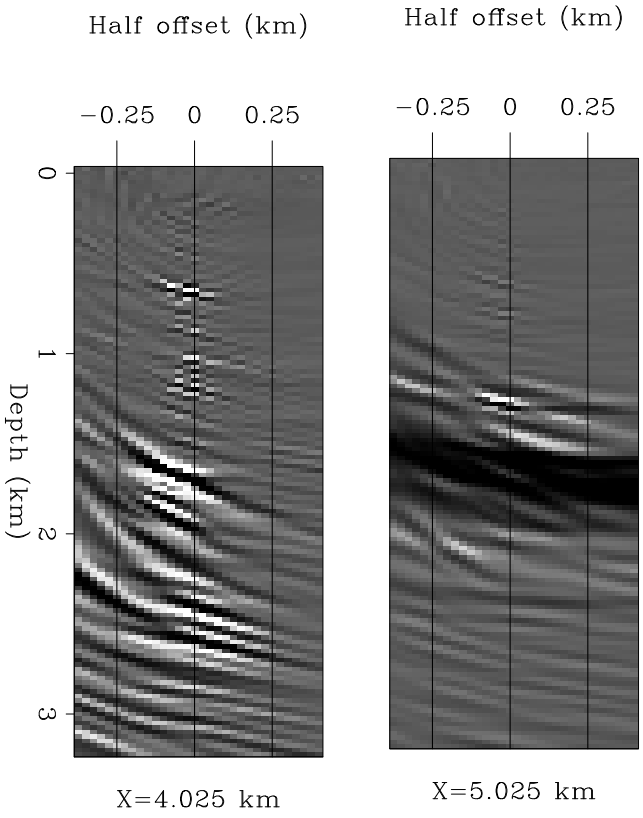


Figure 2: HOCIGs extracted from the prestack image cube. Notice the blurring in the right panel at the depth of the salt edge. biondo1-Cig-all-vz [NR]

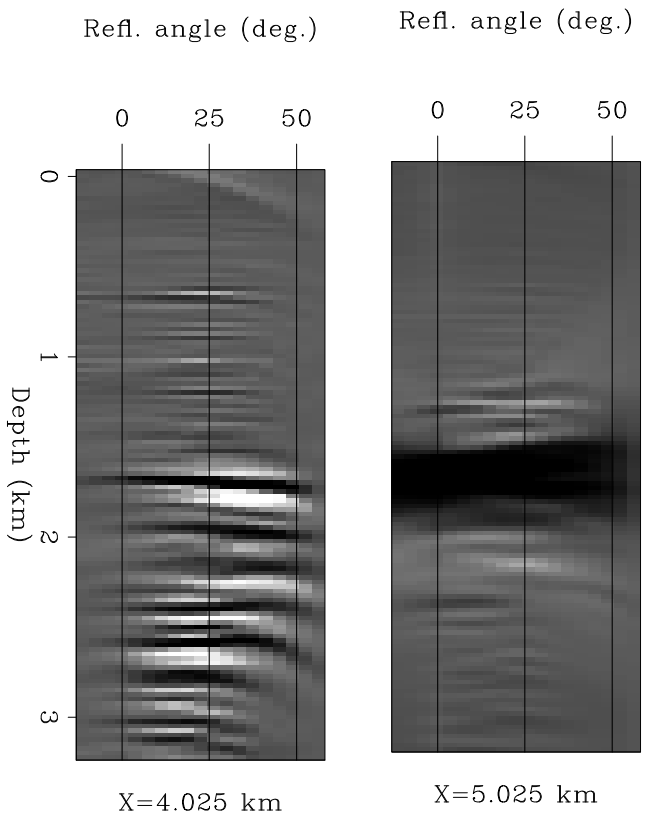


Figure 3: Angle-domain CIGs corresponding to the HOCIGs shown in Figure 2. Notice the blurring in the right panel at the depth of the salt edge. biondo1-Ang-Cig-all-vz [NR]

reflections from the salt edge are focused around zero offset, while the reflections from the flattish reflectors are unfocused. Similarly, after transformation to angle domain (Figure 6) the reflections from the salt edge show a slight moveout that could be used to update the migration velocity function. This task would be impossible if we had to rely solely on the information contained in the ADCIG obtained from HOCIG (Figure 3). However, while the moveouts from the salt edge are clearly interpretable in (Figure 6), the moveouts from flat reflectors are not.

In summary, neither set of CIGs has readily available the information that is needed for velocity updating. In the next section we present a simple method to merge the HOCIG with the VOCIG, and produce a single set of CIGs that satisfies our requirements.

TRANSFORMATION TO DIP-DEPENDENT COMMON IMAGE GATHERS

Two related observations are at the basis of the proposed method. The first one is that the HOCIG and the VOCIG are just particular cases of offset-domain gathers. In general, the offset can be oriented along any arbitrary direction. The second one, is that the offset direction aligned with the apparent geological dip of the imaged event has the unique property of affording the sharpest focusing of the event. The goal of our method is to transform both HOCIGs and VOCIGs into an equivalent set of CIGs (DDOCIGs), for which the effective offset is aligned with the local apparent dips. After the transformation, the DDOCIGs obtained

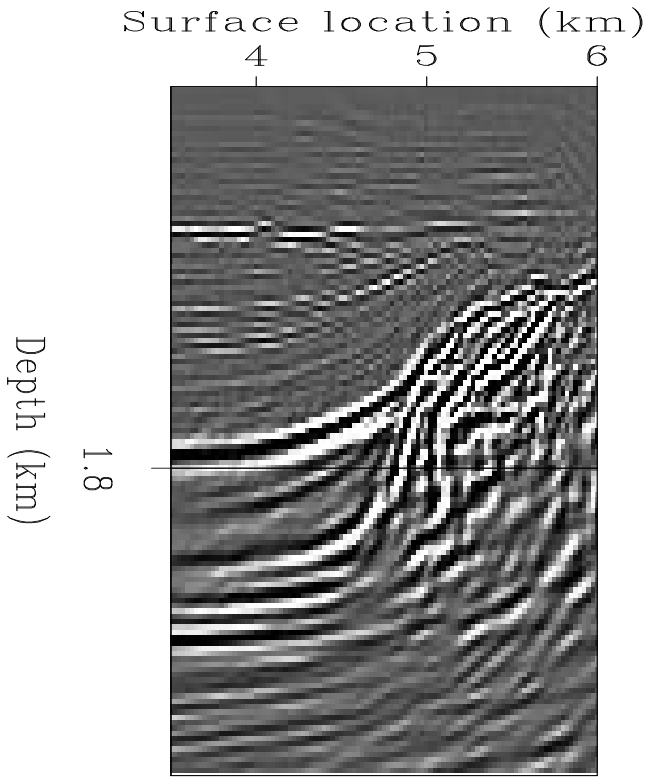


Figure 4: Image of the North Sea data set. The black line superimposed onto the image indicates the depth of the VOCIGs shown in Figure 5 biondo1-image-cig-hz-new [NR]

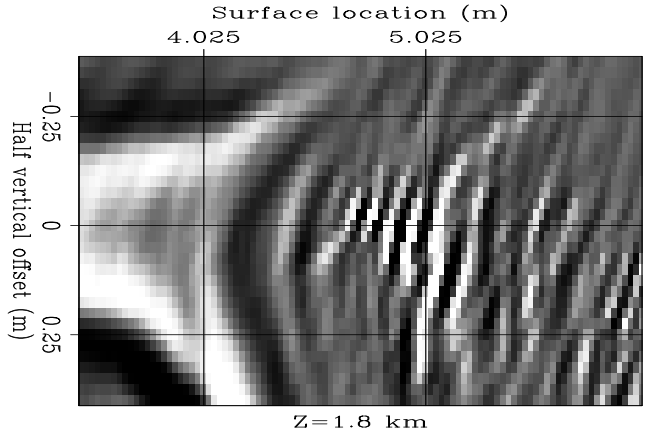


Figure 5: VOCIG extracted from the prestack image cube. Notice the good focus at the horizontal location of the salt edge. biondo1-Cig-1.8-vz-hz [NR]

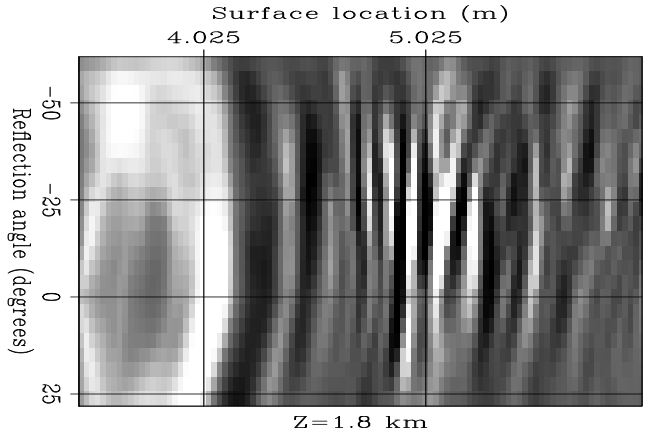


Figure 6: Angle-domain CIG corresponding to the VOCIGs shown in Figure 5. Notice the slight moveout of the the salt edge reflection. biondo1-Ang-Cig-1.8-vz-hz [NR]

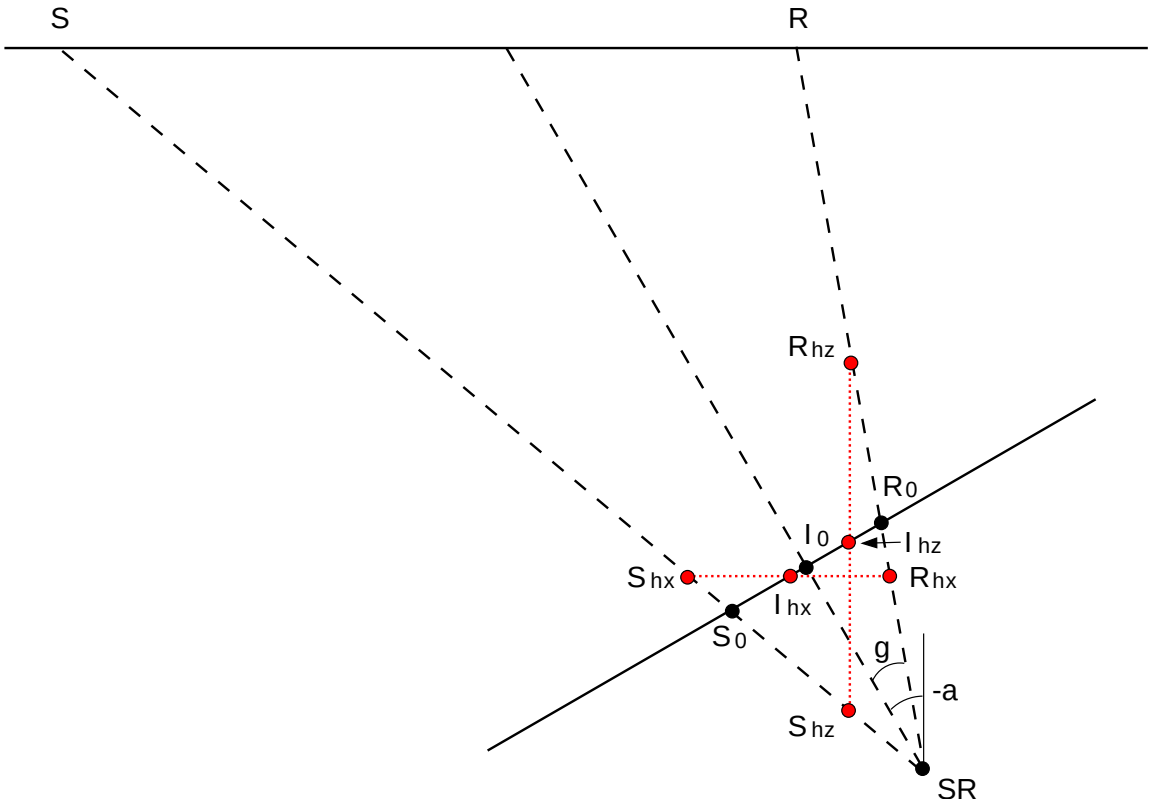


Figure 7: Geometry of the offset-domain CIG for a single event. The migration velocity is assumed to be lower than the true velocity, and thus the events are focused too shallow and above the rays crossing point (SR). biondo1-cig-gen-v4 [NR]

from the HOCIGs and the VOCIGs can be appropriately averaged to obtain a single set of DDOCIGs that contain accurate information for all the geological dips.

Figure 7 illustrates the geometry of the offset-domain CIGs for a single event recorded at the surface for the source location S and receiver location R. The crucial assumption of our geometric construction is that the traveltime along the source ray summed with the traveltime along the receiver ray is the same for all the offset directions and equal to the recording time of the event $\left(|S-S_0|+|R-R_0| = \left|S-S_{x_h}\right|+\left|R-R_{x_h}\right| = \left|S-S_{z_h}\right|+\left|R-R_{z_h}\right|\right)$.

In this sketch, the migration velocity is assumed to be lower than the true velocity, and thus the reflections are imaged too shallow and above the point where the source ray crosses the receiver ray (SR). The line passing through SR, and bisecting the angle formed by the source and receiver ray, is oriented at an angle α with respect to the vertical direction. The angle α is the apparent geological dip of the event after imaging. It would correspond to the true geological dip if the migration velocity were correct. Half of the angle formed between the source and receiver ray is the aperture angle γ.

When HOCIGs are computed, the end point of the source ray (S_{x_h}) and the end point of the receiver ray (R_{x_h}) are at the same depth. The imaging point I_{x_h} is in the middle between

S_{x_h} and R_{x_h} and the imaging offset is $x_h = R_{x_h} - S_{x_h}$. Similarly, when VOCIGs are computed, the end point of the source ray (S_{z_h}) and the end point of the receiver ray (R_{z_h}) are at the same horizontal location. The imaging point I_{z_h} is in the middle between S_{z_h} and R_{z_h} and the imaging offset is $z_h = R_{z_h} - S_{z_h}$. When the offset direction is oriented along the apparent geological dip α (what we called the optimally focusing offset direction), the end point of the source ray is S_0 and the end point of the receiver ray is R_0. The imaging point I_0 is in the middle between S_0 and R_0 and the imaging offset is $h_0 = R_0 - S_0$. It is easy to demonstrate that both I_{x_h} and I_{z_h} lie on the line passing through S_0, I_0 and R_0. The demonstration is based on the assumption that $\left|S_{x_h} - S_0\right| = \left|R_{x_h} - R_0\right|$ and $\left|S_{z_h} - S_0\right| = \left|R_{z_h} - R_0\right|$.

The offsets along the different directions are linked by the following simple relationship, which can be readily derived by trigonometry applied to Figure 7; that is,

$$x_h = \frac{h_0}{\cos\alpha}, \tag{1}$$

$$z_h = \frac{h_0}{\sin\alpha}. \tag{2}$$

Also the shift of the imaging points I_{x_h} and I_{z_h} can be easily expressed in terms of the offset h_0 and the angles α and γ as:

$$\Delta I_{x_h} = \left(I_{x_h} - I_0\right) = -h_0 \tan\gamma \tan\alpha, \tag{3}$$

$$\Delta I_{z_h} = \left(I_{z_h} - I_0\right) = h_0 \frac{\tan\gamma}{\tan\alpha}. \tag{4}$$

Notice the dependence of ΔI_{x_h} and ΔI_{z_h} on the aperture angle γ. This dependence causes events with different aperture angles to be imaged at different locations, even if they originated at the same reflecting point in the subsurface. This phenomenon is related to the well known "reflector-point dispersal" in common midpoint gathers. In this context, this dispersal is a consequence of using a wrong imaging velocity, and we will refer to it as image-point dispersal.

The fact that all three imaging points are aligned along the apparent geological dip allows our transformation to remove the image-point dispersal, and it is crucial to the effectiveness of DDOCIGs. In other words, to transform one set of CIGs into another set we just need to transform the offset axis; the image is then automatically shifted along the apparent geological dip by the right amount. Appendix A demonstrates this fact.

The proposed CIG transformation is a simple dip-dependent non-uniform stretching of the the offset-axis according to the relationships in equations (1) and (2). The transformation is easily implemented in the wavenumber (k_z, k_x) domain, by taking advantage of the well known relationship $\tan\alpha = k_x / k_z$.

After both the HOCIGs and the VOCIGs are transformed, they can be merged together. A simple scheme to merge them is a weighted average, where the weights w_{x_h} and w_{z_h} are set to

$$w_{x_h} = \cos^2\alpha, \tag{5}$$

$$w_{z_h} = \sin^2\alpha. \tag{6}$$

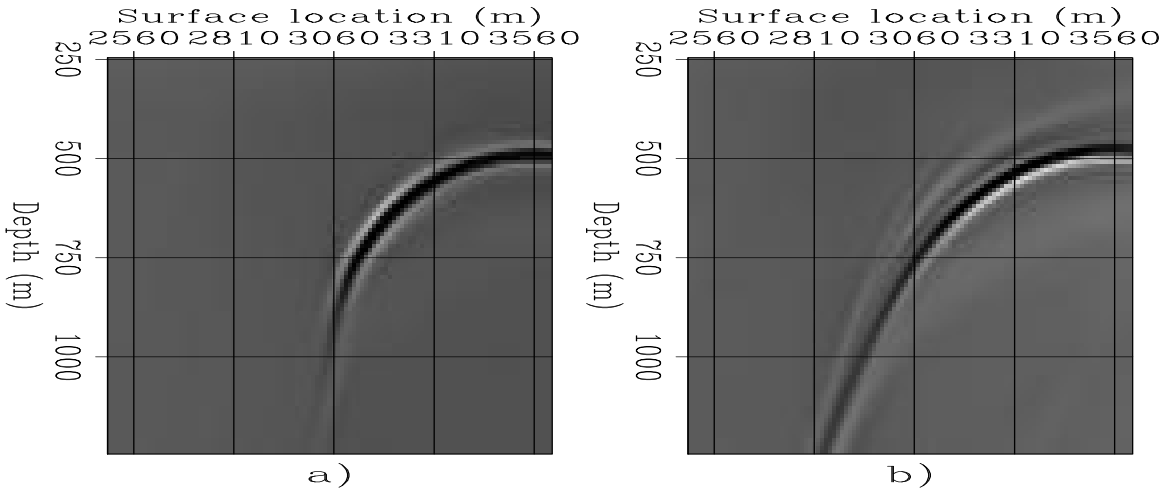


Figure 8: Images of the synthetic data set obtained with a) correct velocity, b) too low velocity by 4%. biondo1-Mig-all-zo-both [CR]

APPLICATION TO A SYNTHETIC DATA SET

To test our proposed method for transforming HOCIGs and VOCIGs into DDOCIGs we modeled and migrated a synthetic data set with a wide range of dips. The reflector has spherical shape with radius of 500 m. The center is at 1,000 meters depth and 3,560 meters horizontal coordinate. The velocity is constant and equal to 2,000 m/s. The data were recorded in 630 shot records. The first shot was located at a surface coordinate of -2,000 meters and the shots were spaced 10 meters apart. The receiver array was configured with asymmetric split-spread geometry. The minimum negative offset was constant and equal to -620 meters. The maximum offset was 4,400 meters for all the shots, with the exception of the first 100 shots (from -2,000 meters to -1,000 meters), where the maximum offset was 5,680 meters in order to record all the useful reflections. To avoid boundary artifacts at the top of the model both sources and receivers were buried 250 meters. Some of the reflections from the top of the sphere were muted out before migration to avoid migration artifacts caused by spurious correlation with the first arrival of the source wavefield.

Figure 8a shows the zero-offset section (stack) of the migrated cubes with the correct velocity (2,000 m/s), and Figure 8b shows the zero-offset section obtained with 4% too low of a velocity (1,920 m/s). Notice that, notwithstanding the large distance between the first shot and the left edge of the sphere (about 5,000 meters), normal incidence reflections illuminate the target only up to about 70 degrees. As we will see in the angle-domain CIGs, the aperture angle coverage shrinks dramatically with the increase of the reflector dip. On the other hand, real data cases are likely to have a vertical velocity gradient that improves the angle coverage of steeply dipping reflectors.

Figures 9 and 10 display sections of the full image cube in the case of the low velocity migration. Figure 9 displays the horizontal-offset image cube, while Figure 10 display the vertical-offset image cube (notice that the offset axis in Figure 10 has been reversed to facil-

itate its visual correlation with the image cube displayed in Figure 9). The side face of the cubes display the CIGs taken at the surface location corresponding to the apparent geological dip of 45 degrees. Notice that the events in the two types of CIGs have comparable shapes, as expected from the geometric analysis presented in the previous section, but their extents are different. The differences between the two image cubes are more apparent when comparing the front faces that show the image at a constant offset of 110 meters (-110 meters in Figure 10). These differences are due to the differences in image-point dispersal for the two offset directions [equation (3) and equation (4)].

Figures 11 and 12 show the image cubes of Figures 9 and 10 after the application of the transformations to DDOCIG, described in equations (1) and (2), respectively. The two transformed cubes are almost identical because both the offset stretching and the image-point dispersal have been removed. The only significant differences are visible in the front face for the reflections corresponding to the top of the sphere. These reflections cannot be fully captured within the vertical-offset image cube because the expression in equation (2) diverges as α goes to zero. Similarly, reflections from steeply dipping events are missing from the horizontal-offset image cube because the expression in equation (1) diverges as α goes to 90 degrees.

The previous figures demonstrate that the proposed transformation converts both HOCIGs and VOCIGs into equivalent DDOCIGs that can be constructively averaged to create a single set of DDOCIGs ready to be analyzed for velocity information. In the following figures, we examine the DDOCIGs obtained by averaging the HOCIGs and VOCIGS using the weights in equations (6), and we compare them with the original HOCIGs and VOCIGs.

We start from analyzing the CIGs obtained when the migration velocity was correct. Figure 13 shows the HOCIGs corresponding to different apparent reflector dips: a) 0 degrees, b) 30 degrees, c) 45 degrees, and d) 60 degrees. The quality of the HOCIGs degrades as dip angle increases. Figure 14 shows the VOCIGs corresponding to the same dips as the panels in Figure 13. In this case, the quality of the VOCIGs improves with the reflector dip.

Figure 15 shows the DDOCIGs corresponding to the same dips as the panels in the previous two figures. Notice that the quality of the DDOCIG is similar to the quality of the HOCIG for small dip angles, and it is similar to the quality of the VOCIG for large dip angles. The focusing of the dipping reflectors (e.g. 60 degrees) is worse than the focusing of the flatter reflectors (e.g. 30 degrees) because of incomplete illumination. In general, the quality of the DDOCIG is "optimal," given the limitations posed by reflector illumination.

The next set of three figures (Figure 16–18) shows the previous offset-domain CIGs transformed into angle domain. The effects of incomplete illumination are more easily identifiable in these gathers than the offset-domain gathers. As for the offset-domain gathers, the angle-domain DDCIGs have consistent quality across the dip range, while the angle-domain gathers obtained from both HOCIG and VOCIG degrade at either end of the dip range.

The next six figures display the same kind of gathers as the past six figures, but obtained when the migration velocity was too low by 4%. They are more interesting than the previous ones, since they are more relevant to velocity updating. Notice that the offset range is doubled

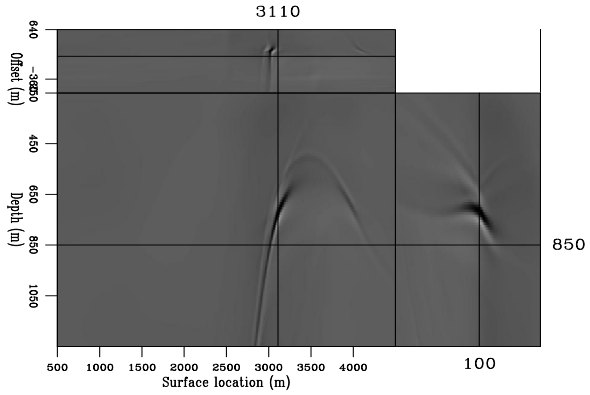


Figure 9: Horizontal-offset image cube when the migration velocity was 4% too low. Notice the differences with the vertical-offset image cube shown in Figure 10. biondo1-Cube-hx-ball-slow-4p [CR]

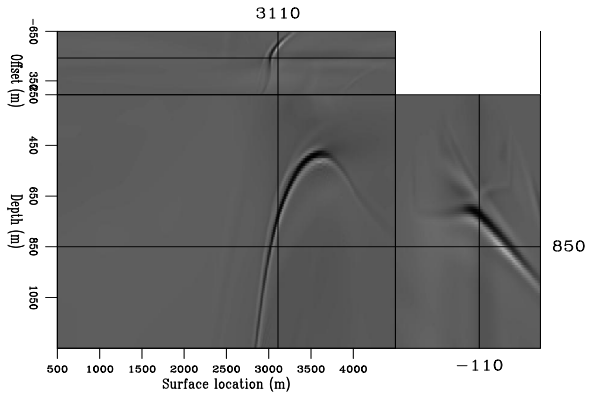


Figure 10: Vertical-offset image cube when the migration velocity was 4% too low. Notice the differences with the horizontal-offset image cube shown in Figure 9. biondo1-Cube-hz-ball-slow-4p [CR]

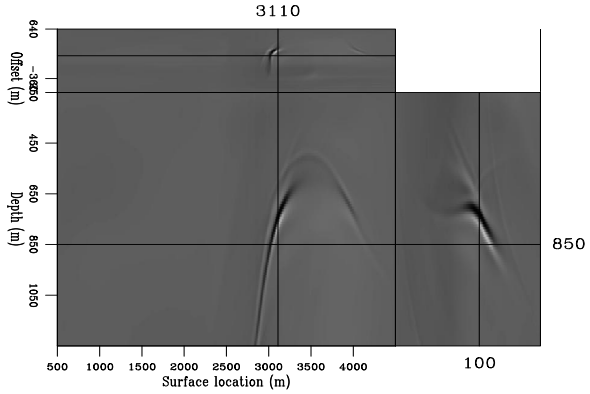


Figure 11: Transformed horizontal-offset image cube. Notice the similarities with the transformed vertical-offset image cube shown in Figure 12. biondo1-Cube-hx-par-ball-slow-4p [CR]

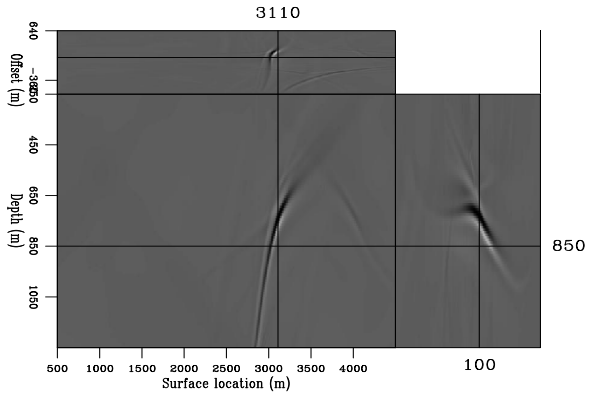


Figure 12: Transformed vertical-offset image cube. Notice the similarities with the transformed horizontal-offset image cube shown in Figure 11. biondo1-Cube-hz-par-ball-slow-4p [CR]

with respect to the previous figures (from $\mp$320 meters to $\mp$640 m) in the attempt to capture within the image cubes all the events, even the ones imaged far from zero offset. For shot profile migration, making the offset range wider is not a trivial additional computational cost.

Figure 19 shows the HOCIGs. The 60 degrees CIG [panel d)] is dominated by artifacts and the corresponding angle-domain CIG shown in Figure 22d would be of difficult use for residual velocity analysis. Figure 20 shows the VOCIGs. As before, the CIGs corresponding to the milder dips are defocused (the artifacts on the left of the panels are caused by the top boundary). The 60 degrees CIG [panel d)] is better behaved than the corresponding HOCIG (Figure 19d), but it is still affected by the incomplete illumination. The DDOCIGs (Figure 21) are the best focused CIGs. Finally the comparison of all the angle-domain CIGs (Figures 22–24) confirm that the DDOCIGS provide the highest resolution and the least-artifact prone ADCIGs, and thus they are the best suited to residual moveout analysis.

CONCLUSIONS

We have introduced a novel transformation of offset-domain Common Image Gathers (CIGs) that applied to either horizontal-offset CIGs (HOCIGs) or vertical-offset CIGs (VOCIGs) transforms them into the equivalent CIGs that would have been computed if the offset direction were aligned along the local geological dip (DDOCIGs). Transformation to DDOCIGs improves the quality of CIGs for steeply dipping reflections by correcting the image cubes from the image-point dispersal. It is particularly useful for velocity analysis when events are not focused around zero offset. The creation of DDOCIGs enables the constructive averaging of HOCIGs with VOCIGs to form DDOCIGs that contain accurate information for all the geological dips. The angle-domain CIGs obtained from the DDOCIGs should provide the best residual moveout information for velocity updating.

We tested the method on a synthetic data set that contains a wide range of dips. The results confirm the theoretical predictions and demonstrate the improvements that are achievable by applying the transformation to DDOCIGs for reflections from steeply dipping reflectors.

ACKNOWLEDGMENTS

We would like to thank Guojian Shan for helping in the computations of the examples of HOCIG and VOCIG from the North Sea data set.

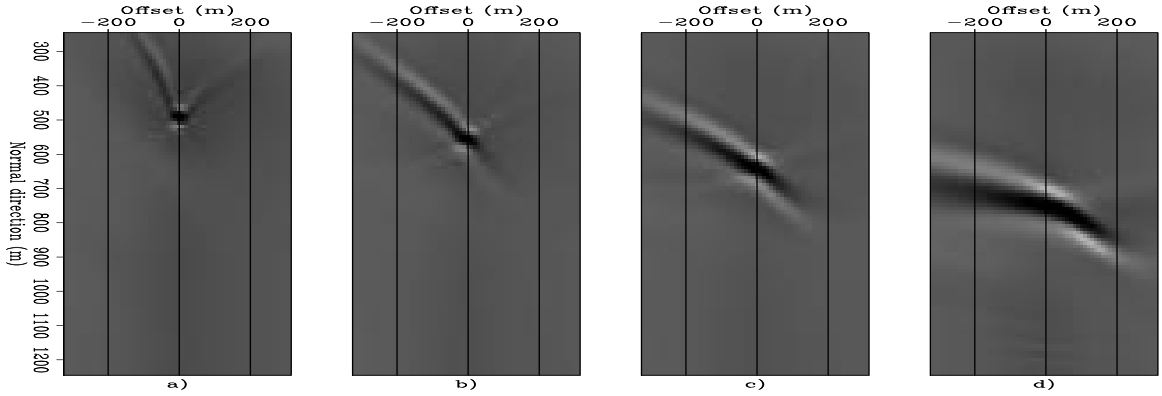


Figure 13: HOCIGs corresponding to different apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was correct. biondo1-Cig-hx-ball-allang [CR]

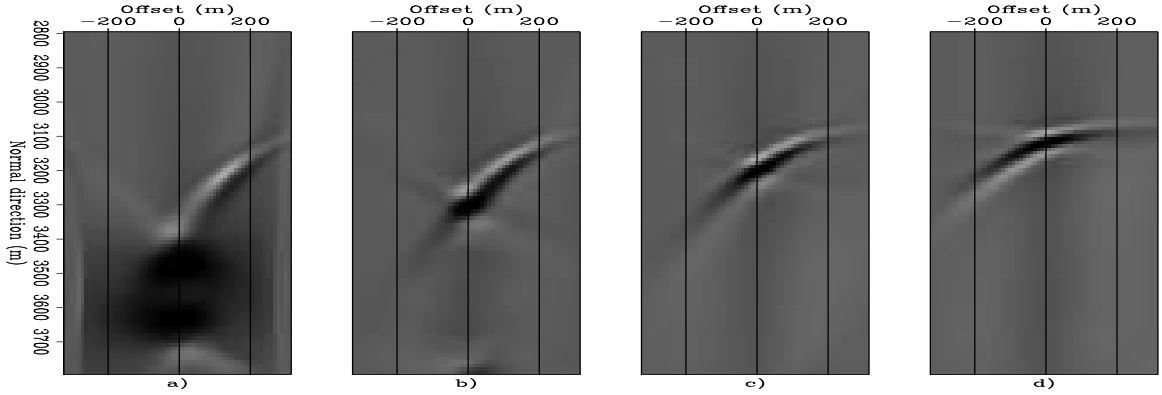


Figure 14: VOCIGs corresponding to different apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was correct. biondo1-Cig-hz-ball-allang [CR]

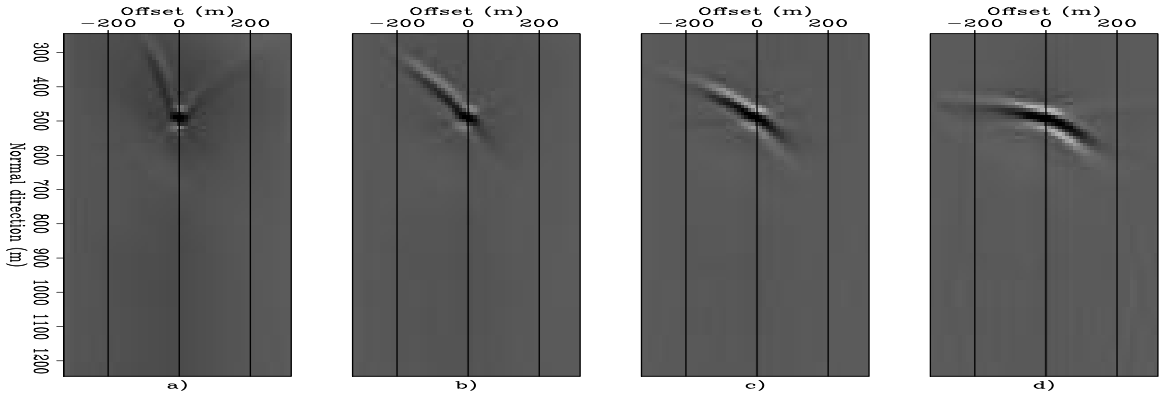


Figure 15: DDOCIGs obtained with the proposed method. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was correct. biondo1-Cig-hrot-ball-allang [CR]

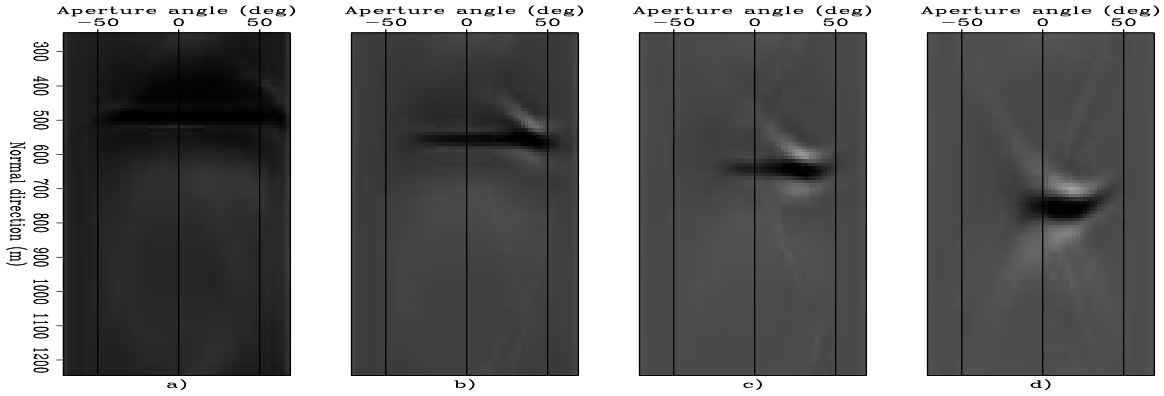


Figure 16: Angle-domain CIGS obtained from the HOCIGs in Figure 13. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was correct. biondo1-Ang-Cig-hx-ball-allang [CR]

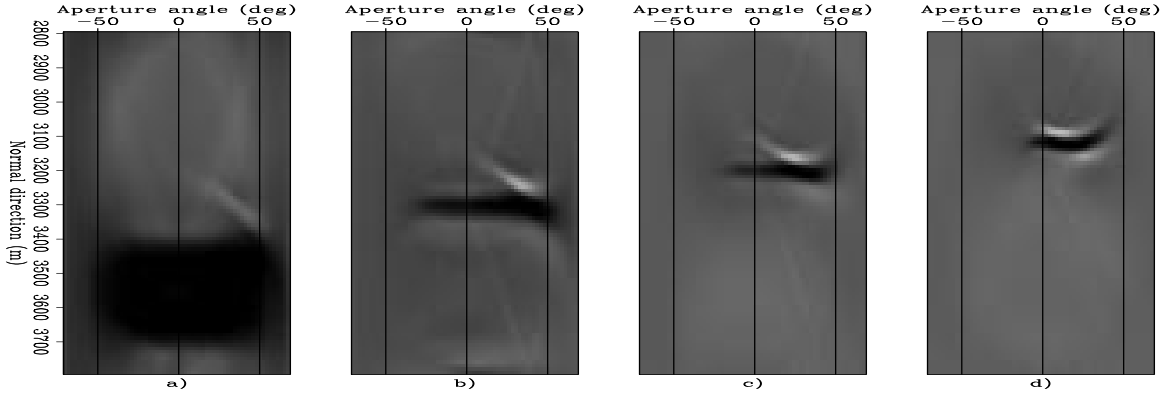


Figure 17: Angle-domain CIGS obtained from the VOCIGs in Figure 14. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was correct. biondo1-Ang-Cig-hz-ball-allang [CR]

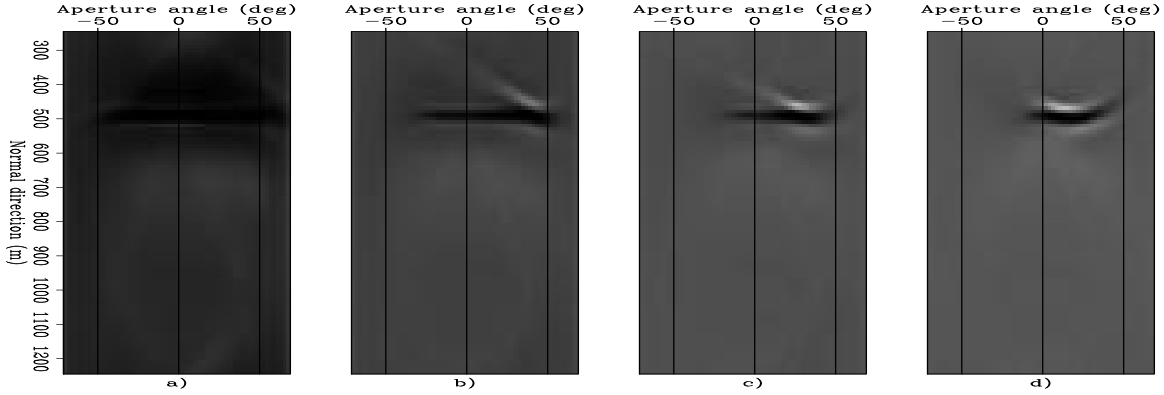


Figure 18: Angle-domain CIGS obtained from the DDOCIGs in Figure 15. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was correct. biondo1-Ang-Cig-hrot-ball-allang [CR]

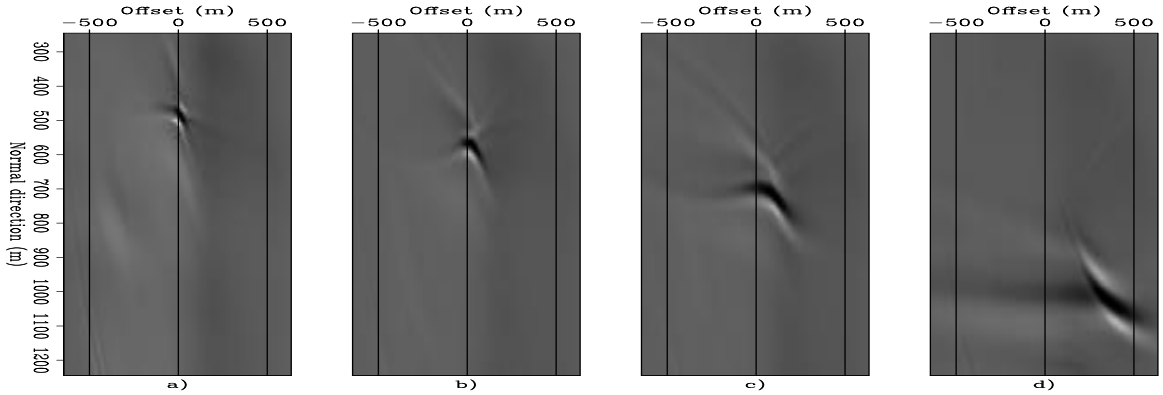


Figure 19: HOCIGs corresponding to different apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was 4% too low. biondo1-Cig-hx-ball-allang-slow-4p [CR]

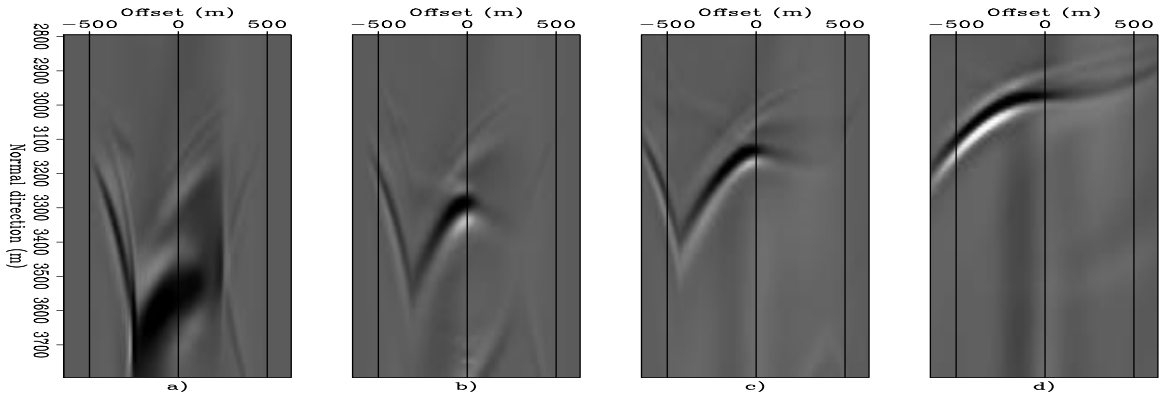


Figure 20: VOCIGs corresponding to different apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was 4% too low. biondo1-Cig-hz-ball-allang-slow-4p [CR]

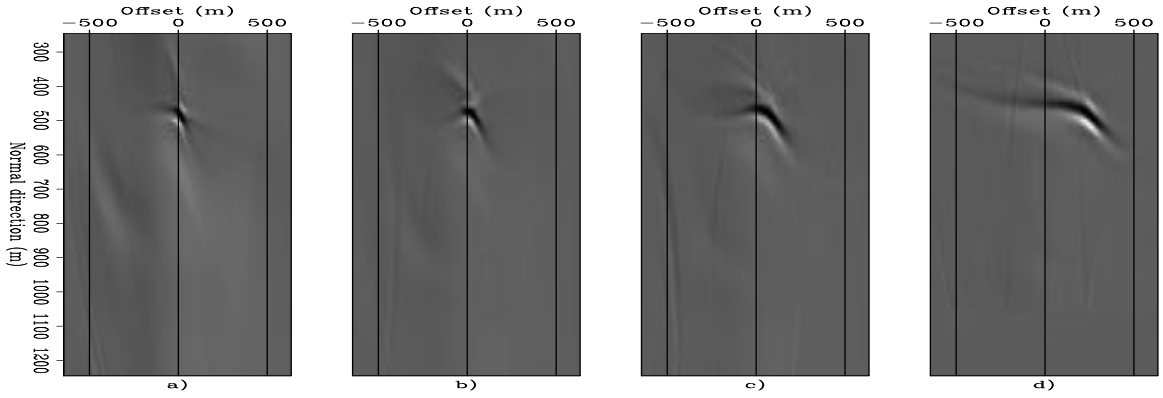


Figure 21: DDOCIGs obtained with the proposed method. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was 4% too low. biondo1-Cig-hrot-ball-allang-slow-4p [CR]

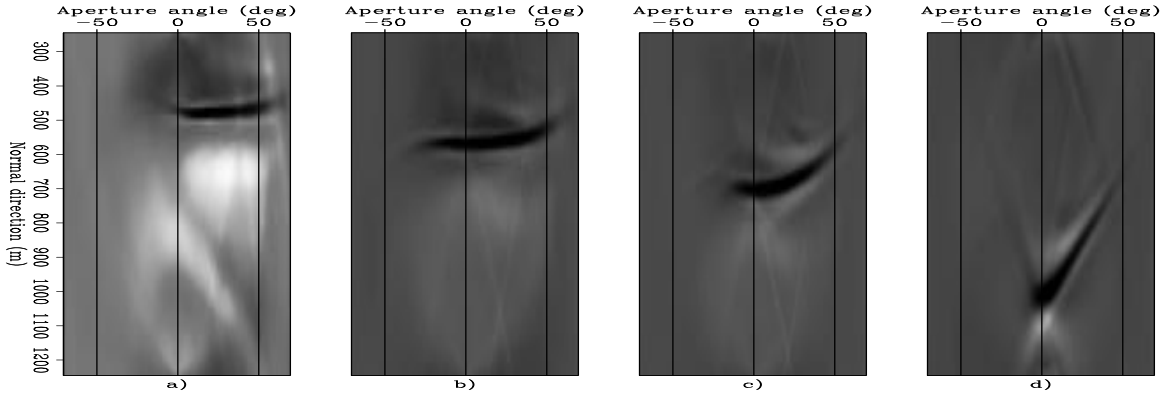


Figure 22: Angle-domain CIGS obtained from the HOCIGs in Figure 13. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was 4% too low. biondo1-Ang-Cig-hx-ball-allang-slow-4p [CR]

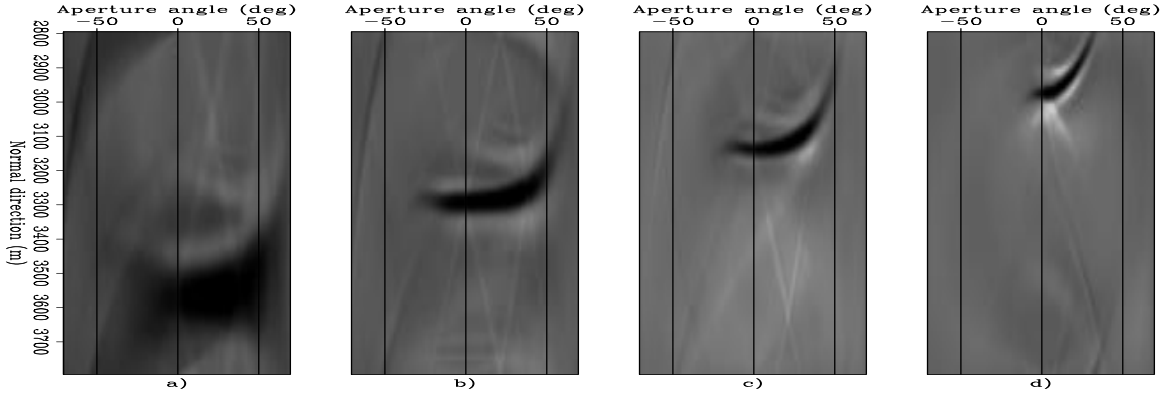


Figure 23: Angle-domain CIGS obtained from the VOCIGs in Figure 14. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was 4% too low. biondo1-Ang-Cig-hz-ball-allang-slow-4p [CR]

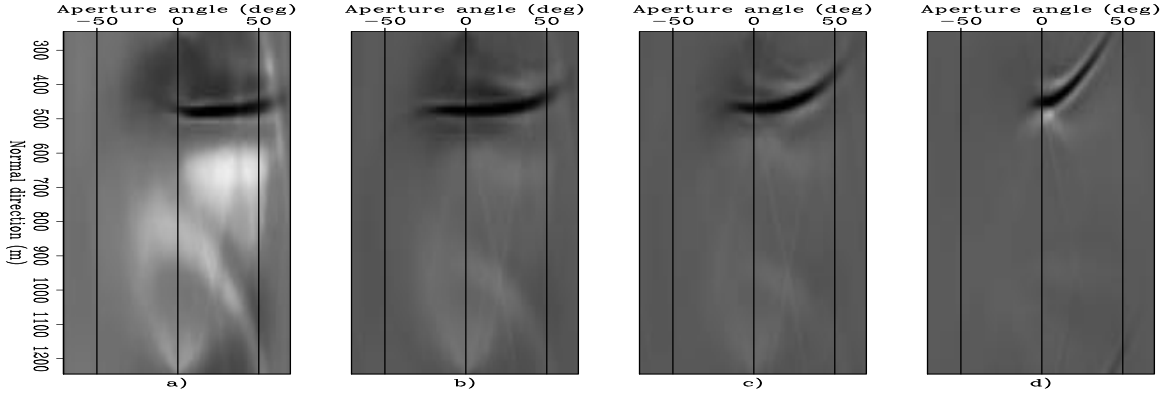


Figure 24: Angle-domain CIGS obtained from the DDOCIGs in Figure 15. The panels correspond to the following apparent reflector dips: a) 0°, b) 30°, c) 45°, and d) 60°. The migration velocity was 4% too low. biondo1-Ang-Cig-hrot-ball-allang-slow-4p [CR]

APPENDIX A

PROOF THAT THE TRANSFORMATION TO DIP-DEPENDENT OFFSET COMMON IMAGE GATHERS (DDOCIG) CORRECTS FOR THE IMAGE-POINT DISPERSAL

This appendix proves that by applying the offset transformations described in equations (1) and (2) we automatically remove the image-point dispersal characterized by equations (3) and (4). The demonstration for the VOCIGs transformation is similar to the one for the HOCIGs transformation, and thus we present only the demonstration for the HOCIGs. HOCIGs are transformed into DDOCIGs by applying the following change of variable of the offset axis x_h, in the vertical wavenumber k_z and horizontal wavenumber k_x domain:

$$x_h = \frac{h_0}{\cos\alpha} = \text{sign}\,(\tan\alpha)\, h_0 \sqrt{1+\tan^2\alpha} = \text{sign}\left(\frac{k_x}{k_z}\right) h_0 \left(1+\frac{k_x^2}{k_z^2}\right)^{\frac{1}{2}}. \tag{A-1}$$

For the sake of simplicity, in the rest of the appendix we will drop the sign in front of expression (A-1) and consider only the positive values of k_x/k_z.

We want to prove that applying (A-1) we also automatically shifts the image by

$$\Delta_z I_{x_h} = h_0 \tan\gamma \tan\alpha \sin\alpha \tag{A-2}$$

in the vertical direction, and

$$\Delta_x I_{x_h} = -h_0 \tan\gamma \tan\alpha \cos\alpha \tag{A-3}$$

in the horizontal direction.

The demonstration is carried out into two steps: 1) we compute the kinematics of the impulse response of transformation (A-1) by a stationary-phase approximation of the inverse Fourier transform along k_z and k_x, 2) we evaluate the dips of the impulse response, relate them to the angles α and γ, and then demonstrate that relations (A-3) and (A-2) are satisfied.

Evaluation of the impulse response of the transformation to DDOCIGs

The transformation to DDHOCIG of an image $I_{x_h}(k_z,k_x,x_h)$ is defined as

$$I_0(k_z,k_x,k_h) = \int dh_0 I_0(k_z,k_x,h_0) e^{ik_h h_0} = \int dx_h \left(\frac{dh_0}{dx_h}\right) I_{x_h}(k_z,k_x,x_h) e^{ik_h x_h \left(1+\frac{k_x^2}{k_z^2}\right)^{-\frac{1}{2}}}. \tag{A-4}$$

The transformation to DDHOCIG of an impulse located at $(\bar{z},\bar{x},\bar{x_h})$ is thus (after inverse Fourier transforms):

$$\widetilde{\text{Imp}}(z,x,h_0) = \int dk_h \int dx_h \int dk_x \int dk_z \left(\frac{dh_0}{dx_h}\right) e^{i\left\{k_h\left[\bar{x_h}\left(1+\frac{k_x^2}{k_z^2}\right)^{-\frac{1}{2}}-h_0\right]+k_z(\bar{z}-z)+k_x(\bar{x}-x)\right\}}. \tag{A-5}$$

We now approximate by stationary phase the inner double integral. The phase of this integral is,

$$\Phi \equiv k_h \left[\bar{x_h} \left(1 + \frac{k_x^2}{k_z^2} \right)^{-\frac{1}{2}} - h_0 \right] + k_z (\bar{z} - z) + k_x (\bar{x} - x) \tag{A-6}$$

The stationary path is defined by the solutions of the following system of equations:

$$\frac{\partial \Phi}{\partial k_z} = k_h \bar{x_h} \frac{k_x^2}{k_z^3} \left(1 + \frac{k_x^2}{k_z^2} \right)^{-\frac{3}{2}} + (\bar{z} - z) = 0, \tag{A-7}$$

$$\frac{\partial \Phi}{\partial k_x} = -k_h \bar{x_h} \frac{k_x}{k_z^2} \left(1 + \frac{k_x^2}{k_z^2} \right)^{-\frac{3}{2}} + (\bar{x} - x) = 0, \tag{A-8}$$

By moving both $(\bar{z} - z)$ and $(\bar{x} - x)$ on the right of equations (A-7) and (A-8), and then dividing equation (A-7) by equation (A-8), we obtain the following relationships between $(\bar{z} - z)$ and $(\bar{x} - x)$:

$$\frac{\bar{z} - z}{\bar{x} - x} = -\frac{k_x}{k_z}. \tag{A-9}$$

Furthermore, by multiplying equations (A-7) by k_z and equation (A-8) by k_x, and then substituting them appropriately in the phase function (A-6), we can evaluate the phase function along the stationary path as

$$\Phi_{\text{stat}} = k_h \left[\bar{x_h} \left(1 + \frac{k_x^2}{k_z^2} \right)^{-\frac{1}{2}} - h_0 \right], \tag{A-10}$$

that becomes, by substituting equation (A-9),

$$\Phi_{\text{stat}} = k_h \left\{ -\bar{x_h} \left[1 + \frac{(\bar{z} - z)^2}{(\bar{x} - x)^2} \right]^{-\frac{1}{2}} - h_0 \right\}. \tag{A-11}$$

Notice that the minus sign comes from the sign function in expression (A-1). By substituting expression (A-11) in equation (A-5) it is immediate to evaluate the kinematics of the impulse response as

$$h_0 = -x_h \left[1 + \frac{(\bar{z} - z)^2}{(\bar{x} - x)^2} \right]^{-\frac{1}{2}} \tag{A-12}$$

Evaluation of the image shift as a function of α ad γ

The final step is to take the derivative of the impulse response of equation (A-12) and use the relationships of these derivatives with $\tan\alpha$ and $\tan\gamma$.

$$\frac{\partial z}{\partial x} = \tan\alpha = \sqrt{\frac{x_h^2}{h_0^2} - 1}, \tag{A-13}$$

$$\frac{\partial z}{\partial x_h} = \tan\gamma = (\bar{x}-x)\frac{\frac{x_h}{h_0}}{\sqrt{\frac{x_h^2}{h_0^2}-1}} = (\bar{z}-z)\frac{\frac{x_h}{h_0}}{\frac{x_h^2}{h_0^2}-1}. \tag{A-14}$$

Substituting equations (A-13) and (A-14) into

$$\Delta_z I_{x_h} = \bar{z}-z = h_0 \tan\gamma \tan\alpha \sin\alpha \tag{A-15}$$

$$\Delta_x I_{x_h} = \bar{x}-x = -h_0 \tan\gamma \tan\alpha \cos\alpha. \tag{A-16}$$

.

and after some algebraic manipulations we prove the thesis.

REFERENCES

Biondi, B., and Shan, G., 2002, Prestack imaging of overturned reflections by reverse time migration: 72nd Ann. Internat. Meeting, Soc. of Expl. Geophys., Expanded Abstracts, to be published.

Biondi, B., 2002, Equivalence of source-receiver migration and shot-profile migration: SEP–**112**, 109–116.

Clapp, R., and Biondi, B., 2000, Tau domain migration velocity analysis using angle CRP gathers and geologic constrains: 70th Ann. Internat. Mtg., Soc. Expl. Geophys., 926–929.

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

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

Rickett, J., and Sava, P., 2001, Offset and angle domain common-image gathers for shot-profile migration: 71th 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. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 296–299.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 83–96

Deconvolution imaging condition for reverse-time migration

Alejandro A. Valenciano and Biondo Biondi[1]

ABSTRACT

The reverse-time migration imaging condition can be improved by computing the reflection strength at each subsurface point as the zero lag value of the deconvolution of the receiver wavefield by the source wavefield. I show that by using this approach it is possible to eliminate image artifacts due to wavefield multipathing through velocity anomalies. I also show that it has the advantage of handling better amplitudes during imaging.

INTRODUCTION

The most common implementation of shot-profile reverse-time migration (Etgen, 1986; Biondi, 2002) uses the zero lag of the cross-correlation of the source and the receiver wavefields as the imaging condition. This implementation has the advantage of being robust and honoring the kinematics of Claerbout's imaging principle but does not honor the dynamics of the problem, which results in the loss of resolution and amplitude accuracy (Claerbout, 1971).

Another drawback of this imaging condition is that it creates image artifacts when there is a complex propagation pattern, e.g., a low velocity anomaly that cause wavefield multipathing. Let us consider waves propagating in a homogeneous medium with a velocity anomaly and a flat reflector. After the wave traveling directly from the shot to the reflector arrives, a second wave arrives, which has traveled along a different path due to the low velocity anomaly. If the second wave is not accounted for in the imaging process, the single reflector will be imaged as more than one reflector. This could mislead the geological interpretation.

In this paper, we introduce an imaging condition that computes the reflection strength as the zero lag of the deconvolution of the receiver wavefield by the source wavefield. This new process can account for the second wave arrival in the imaging. We implemented two equivalent deconvolution methods: one in the time domain based on least-squares inversion filtering and the other in the Fourier domain.

We illustrate the wavefield deconvolution imaging condition with two different data sets. One created by the convolution of a minimum-phase, band-limited wavelet with a spike series, and the other by wave equation modeling and downward propagation.

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

REFLECTOR MAPPING IMAGING CONDITION AND IMAGE ARTIFACTS

Claerbout (1971) expresses the reflector mapping imaging condition as follows:

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

where x is the horizontal coordinate, z is the depth, and t_d is the time at which the source wavefield $\mathbf{d}(x,z,t_d)$ and the receiver wavefield $\mathbf{u}(x,z,t_d)$ coincide in time and space. This principle states that the reflectivity strength $\mathbf{r}(x,z)$ depends only on the source wavefield and on the receiver wavefield at time t_d.

A practical way to compute the reflectivity strength is discussed in Claerbout's paper (Claerbout, 1971). The reflectivity strength is computed as:

$$\mathbf{r}(x,z) = (\mathbf{u} \star \mathbf{d})(x,z,\tau = 0), \tag{2}$$

where $\star$ means cross-correlation and τ is the lag. This is commonly used in the industry. It has the advantage of being robust, but has the disadvantage of not computing the correct amplitudes (Claerbout, 1971).

A more general imaging condition can be stated, computing the reflectivity strength as:

$$\mathbf{r}(x,z) = \frac{\mathbf{u}}{\mathbf{d}}(x,z,\tau = 0), \tag{3}$$

where the division means deconvolution in time of the receiver wavefield by the source wavefield for each (x,z) and τ is the lag. It has the potential of accounting for wavefield multipathing during imaging, thus avoiding the creation of image artifacts in the presence of velocity anomalies.

Figures 1 to 3 show the comparison of wavefield deconvolution with wavefield cross-correlation imaging condition. The first row, in Figure 1, simulates the two wavefields coinciding at the reflector depth. The result of the cross-correlation and the result of the deconvolution is shown in the second row. For each case, the zero lag of the wavefield cross-correlation or the zero lag of the wavefield deconvolution is assigned as the reflectivity strength at this depth.

The first row in Figure 2 / 3 simulates the two wavefields at a deeper / shallower depth than the reflector depth. The second row shows the result of the cross-correlation and the result of the deconvolution. The zero lag value of the wavefield cross-correlation has a value different than zero, thus creates an image artifact at a deeper / shallower depth. In the case of deconvolution imaging condition, the zero lag value is zero, thus no image artifacts are created.

The imaging condition stated in equation (3) makes the strong assumption that the receiver wavefield $\mathbf{u}(x,z,t)$ can be computed by convolving the source wavefield $\mathbf{d}(x,z,t)$ by the reflectivity strength $\mathbf{r}(x,z)$. As we will discuss later, this is true at the reflector depth, but might not be true at a different depth.

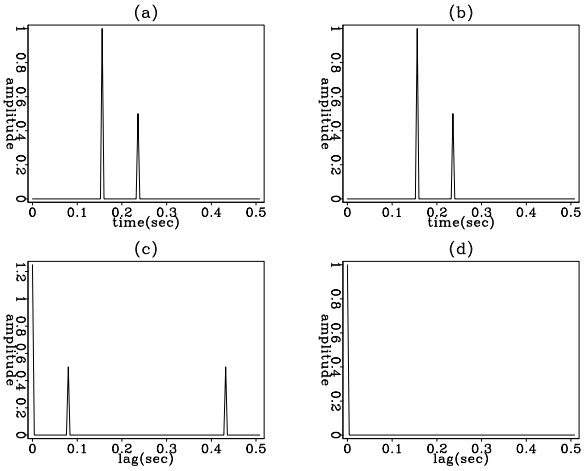


Figure 1: Wavefields coinciding at the reflector depth. (a) Source wavefield. (b) Receiver wavefield. (c) Wavefields cross-correlation. (d) Wavefields deconvolution. alejandro1-spike [ER]

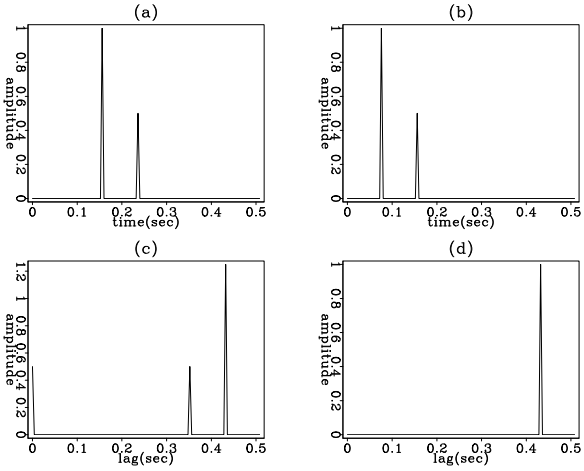


Figure 2: Wavefields at a depth deeper the reflector depth. (a) Source wavefield. (b) Receiver wavefield. (c) Wavefields cross-correlation. (d) Wavefields deconvolution. alejandro1-spike1 [ER]

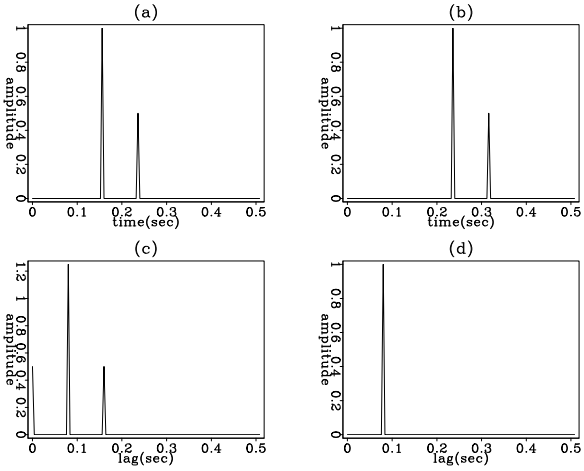


Figure 3: Wavefields at a depth shallower the reflector depth. (a) Source wavefield. (b) Receiver wavefield. (c) Wavefields cross-correlation. (d) Wavefields deconvolution. alejandro1-spike2 [ER]

DECONVOLUTION IN THE TIME DOMAIN

Deconvolution in the time domain can be implemented in terms of the following fitting goal for each (x,z) location:

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

where $\mathbf{D}$ is a convolution matrix whose columns are downshifted versions of the source wavefield $\mathbf{d}$.

The least-squares solution of this problem is

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

where $\mathbf{D}'$ is the adjoint of $\mathbf{D}$. A damped solution may be 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 ε is a small positive number. Equation (6) can be written in terms of the fitting goals

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

where $\mathbf{I}$ is the identity matrix. This approach can be computational efficient if the time window is not too large and we use a Conjugate Gradient as optimization engine. However, it has the disadvantage of relying on a linear inversion process that may or may not converge to the global minimum. A way to overcome this problem, obtaining an analytical solution, is to implement equation (6) in the Fourier domain, as we do in the next section.

DECONVOLUTION IN THE FOURIER DOMAIN

Jacobs (1982) compares various imaging methods for shot-profile migration. He shows that the deconvolution imaging condition

$$\mathbf{r}(\omega) = \frac{\mathbf{u}(\omega)\overline{\mathbf{d}}(\omega)}{\mathbf{d}(\omega)\overline{\mathbf{d}}(\omega) + \varepsilon^2} \tag{8}$$

is stable. The same imaging condition was used by Lee et al. (1991) in split-step migration.

Since we are only interested in the zero-time lag, the reflection strength can be computed as

$$\mathbf{r}(x,z,\tau=0) = \sum_{\omega}^{\omega_{Nyq}} \mathbf{r}(x,z,\omega). \tag{9}$$

where ω_{Nyq} is the Nyquist frequency.

An advantage of working in the Fourier domain is that the problem does not need to be stated as an inversion problem.

DECONVOLUTION IN THE FOURIER DOMAIN WITH BAND-LIMITED DATA

In equation (9) the reflectivity strength is affected by the bandwidth of the data. Let us consider the case where the source and receiver wavefields coincide at the reflector depth and both have a similar frequency content. The best situation we could have is for their division to be a box function. This is unlikely to happen in a real case. Inside the data bandwidth the division is a constant value. But outside of the data bandwidth we may try to divide small numbers by small numbers, which has the potential to be unstable. In equation (8) we use a damping factor to avoid this source of instability but when we apply equation (9) we end up with values that were supposed to be zero contributing to the reflectivity strength.

A different implementation of equation (8), forcing hard zeros when $\mathbf{d}(\omega)\overline{\mathbf{d}}(\omega) < \varepsilon^2$, reduces the impact of band-limited data in the reflectivity strength calculation as

$$\mathbf{r}(\omega) = \begin{cases} \frac{\mathbf{u}(\omega)}{\mathbf{d}(\omega)} & \mathbf{d}(\omega)\overline{\mathbf{d}}(\omega) > \varepsilon^2 \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

There is another source of error for the band-limited. The Fourier pair of the box function is a sinc function. In the extreme case of a infinite wide box the Fourier pair is a delta function centered at zero lag. As the box is getting narrower in the Fourier domain, the delta becomes a wider sinc function in the time domain. Therefore, the reflectivity strength $\mathbf{r}(x, z, \tau = 0)$ is a scaled version of his infinite bandwidth version. We can compensate for the bandwidth of the data by computing the zero lag of the deconvolution as

$$\mathbf{r}(x, z, \tau = 0) = \frac{\omega_{Nyq}}{\Delta\omega_{BW}} \sum_{\omega}^{\omega_{Nyq}} \mathbf{r}(x, z, \omega), \tag{11}$$

where ω_{Nyq} is the Nyquist frequency and $\Delta\omega_{BW}$ is the bandwidth, then zero lag of the deconvolution corresponds to the reflectivity strength. In this case the bandwidth $\Delta\omega_{BW}$ is defined as the frequency range where the inequality $\mathbf{d}(\omega)\overline{\mathbf{d}}(\omega) > \varepsilon^2$ holds.

TEST WITH SYNTHETIC DATA (CONVOLUTIONAL MODEL)

We tested the different implementations of the imaging condition with synthetic data. The data simulates the case when source and receiver wavefields coincide at reflector depth and the case when they coincide at a shallower depth. The data are constructed by convolving a minimum-phase, $25Hz$ central frequency, band-limited wavelet with a series of spikes.

Figure 4 shows the deconvolution of the signal by itself. The signal was constructed by convolving the wavelet with two spikes. This simulates the situation where source and receiver wavefields coincide at reflector depth. As expected, the result is a delta function centered at zero lag with no difference between the two deconvolution methods (Figures 4d and 4e). Figure 4c shows the result of the cross-correlation for the sake of comparison. The cross-correlation result differs from the deconvolution result in resolution, but still is a symmetric function centered at zero lag.

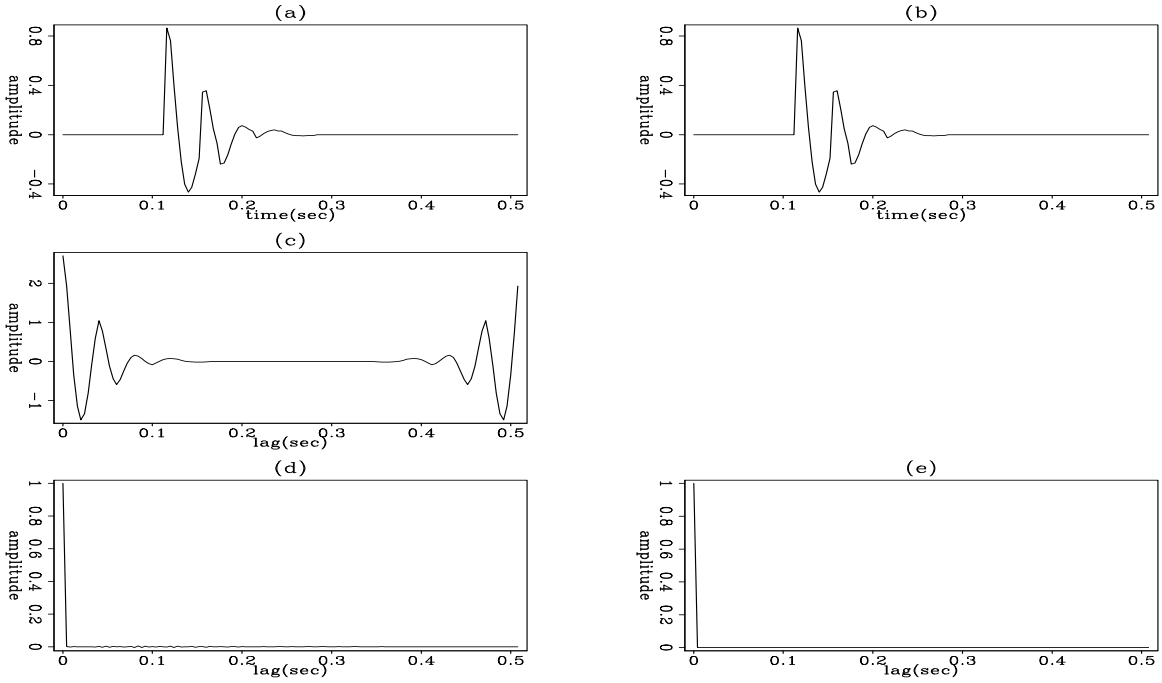


Figure 4: Source and receiver wavefields shifted from zero, simulating wavefields coinciding at the reflector depth. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-shift [ER]

Figure 5 shows the result of deconvolving the same signal shifted to the right (Figure 5b) by the unshifted signal (Figure 5a). This simulates the situation where the receiver and the source wavefield coincide at a depth shallower than the reflector depth. The result is a shifted delta function. No significant differences can be seen between the two convolution methods. In this situation the cross-correlation (Figure 5c) produces an erroneous image since the zero lag is different than zero.

Figure 6 shows the deconvolution of the same signal (in Figure 4b) contaminated by more spikes (Figure 6b) with the original signal (Figure 6a). This resembles a real situation when source and receiver wavefields coincide at reflector depth. The deconvolution method based on least squares inversion in time gets the correct value at zero lag but does not converge to the global minimum. In the Fourier domain the delta at zero is recuperated and some energy comes at the end of the signal due to symmetric boundary conditions. Since we are only interested in the zero lag value, both deconvolution methods could be used. The result of the cross-correlation (Figure 6c) has a maximum at zero lag as expected.

Figure 7 shows the deconvolution of the same signal (in Figure 6b) shifted to the right (Figure 7b). This simulates the situation where the receiver and the source wavefield coincide at a depth shallower than the reflector depth. The deconvolution method based on least squares inversion in the time domain recovers the correct shifted spike. As we saw in the previous case, when some energy exist before the onset of the reflector energy in the receiver wavefield,

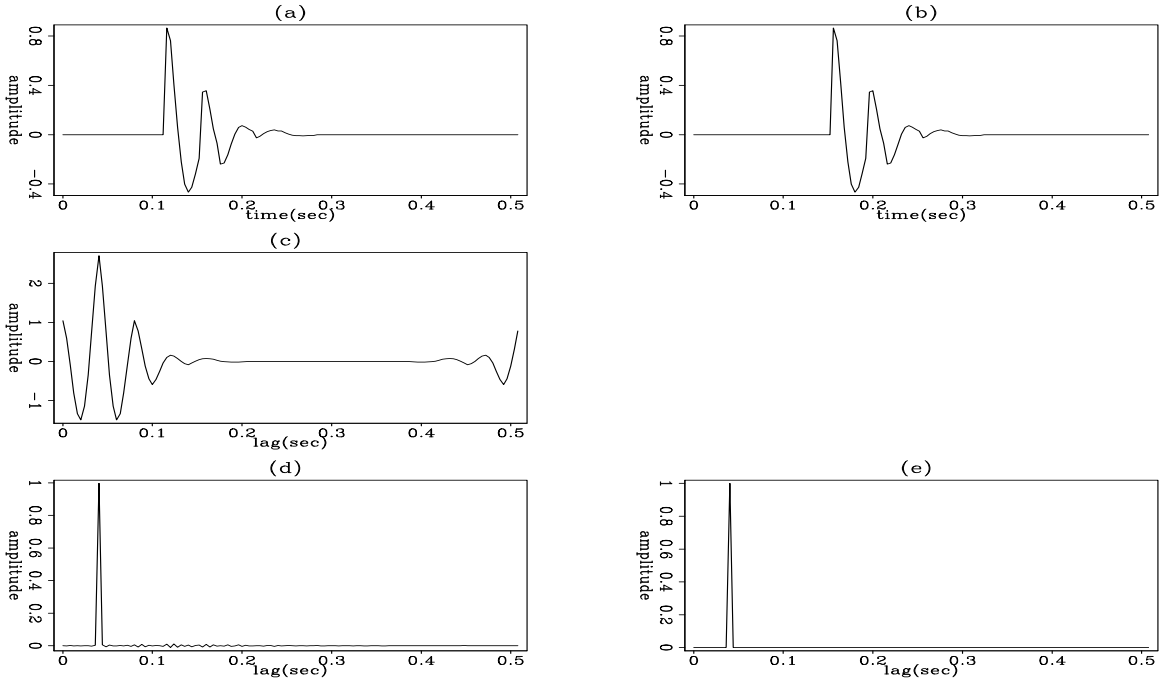


Figure 5: Source and receiver wavefields shifted from zero and each others, simulating wavefields coinciding at a depth shallower the reflector depth. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-shift_dt [ER]

the least squares fails to reach the global minimum. In the Fourier domain, we recover the shifted spike and some energy comes at the end of the signal due to symmetric boundary conditions. In this case the deconvolution in the Fourier domain has a better performance than the deconvolution in the time domain since there is no energy at zero lag, as was theoretically predicted.

TEST WITH SYNTHETIC DATA (WAVE EQUATION)

Now, we test our imaging condition with a more realistic model. The data are modeled with a wave equation-finite differences program. Figure 8 shows the velocity model. Note the low velocity anomaly at 300 m depth. Also, a flat, constant-impedance contrast interface is located at 700 m. The low velocity anomaly creates multipathing that we want to include during the imaging.

From the whole receiver $\mathbf{u}(z,x,t)$ and source $\mathbf{d}(z,x,t)$ wavefields we extracted two constant depth planes: one at the reflector depth (Figure 9) and one shallower (Figure 10). In both figures we can see two reflectors, one due to the direct arrival and the other due to the second wave arrival produced by the low velocity anomaly.

Four traces were extracted for this test at two different offsets (1000 m and 2000 m).

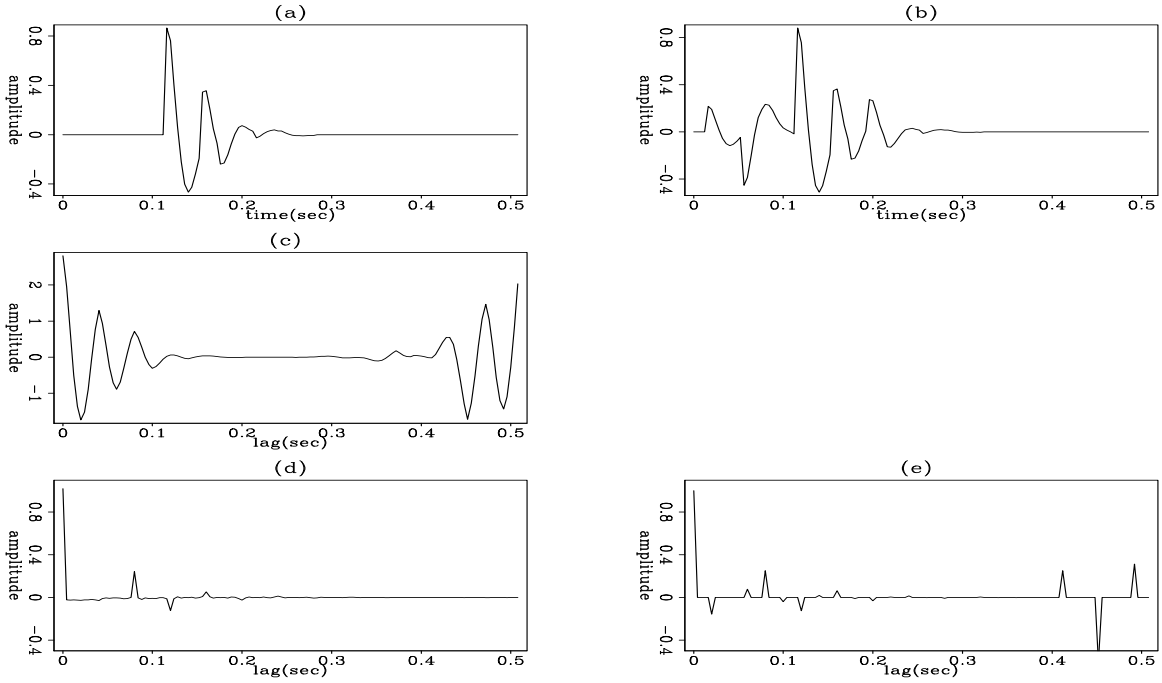


Figure 6: Source and receiver wavefields shifted from zero. It simulates two wavefields coinciding at the reflector depth. Receiver wavefield contaminated with more events. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-shift_mult [ER]

Figures 11a and 12a show the source wavefield and Figures 11b and 12b the receiver wavefield at reflector depth. Figures 13a and 14a show the source wavefield and Figures 13b and 14b the receiver wavefield at a shallower depth.

Figures 11 and 12 show the deconvolution of the receiver by the source wavefield at the reflector depth. As we expected, there is a maximum at zero lag. The result is very similar using both deconvolution methods. Convolution of the source wavefield with the reflectivity seems to be a good modeling operator, at least at the reflector depth, since the data residual in the least squares inversion is small.

Figures 13 and 14 show the deconvolution of the receiver by the source wavefield at a depth shallower than the reflector depth. As we expected, the value at zero lag is an order of magnitude smaller than the value at zero lag at the reflector depth. The result is very similar using both deconvolution methods. For an offset of 2000 m and a depth of 500 m, the data residual in the least squares inversion is small, indicating that the convolution of the source wavefield with the reflectivity fits the receiver wavefield. But for an offset of 1000 m and a depth of 500 m, the data residual in the least squares inversion is not small. The explanation for this result can be found looking at Figures 13a and 13b. We can note that the separation between the two arrivals in the source and receiver wavefields is not the same. Thus the source wavefield convolved with a simple delta shifted from the zero lag (as reflectivity) is not enough to explain the receiver wavefield. The impact of this issue in the final image quality

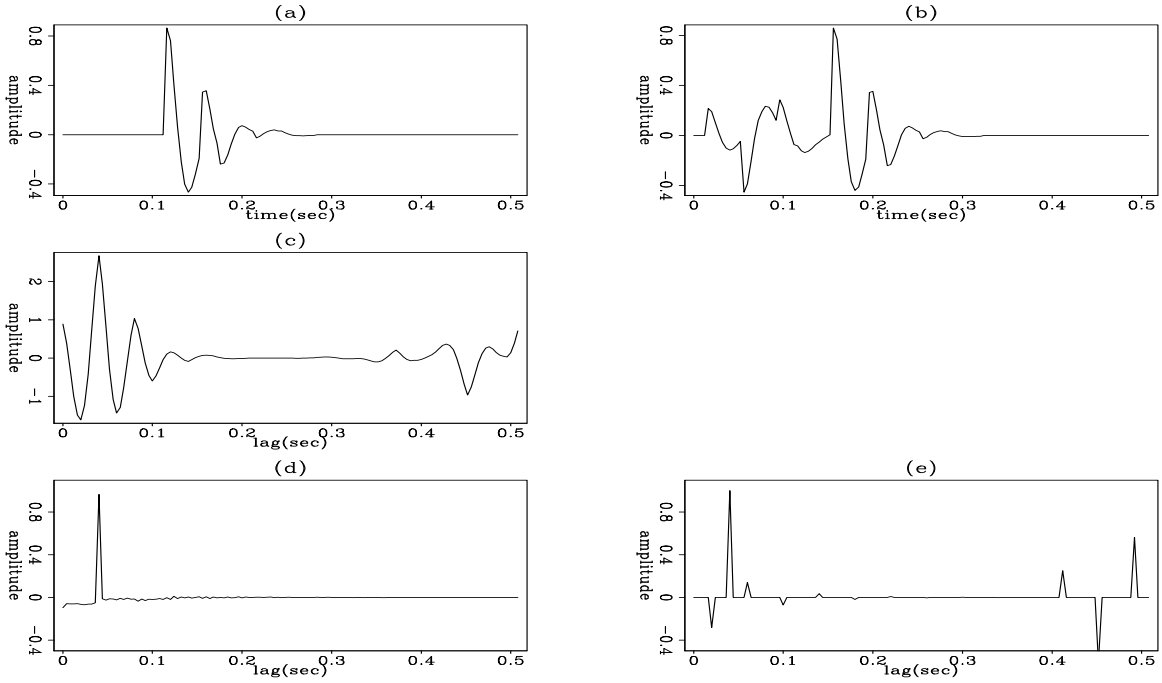


Figure 7: Source and receiver wavefields shifted from zero and each others. It simulates two wavefields coinciding at a depth shallower the reflector depth. Receiver wavefield contaminated with more events. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-shift_mult_dt [ER]

needs further investigation.

Comparing the deconvolution with the cross-correlation imaging condition in Figures 11, 12, 13, and 14, we conclude that deconvolution imaging condition effectively attenuates the image artifacts and handles the amplitudes better. In addition, the deconvolution imaging condition does a better job than the cross-correlation in preserving the amplitudes through the offset.

CONCLUSIONS

The cross-correlation of the receiver wavefield by the source wavefield imaging condition has the advantage of being robust and honoring the kinematics of Claerbout's imaging principle. However, it has the disadvantage of losing resolution and creating image artifacts in the presence of multipathing.

A better imaging condition is obtained by computing the reflection strength as the zero lag value of the deconvolution of the receiver wavefield by the source wavefield. This approach was implemented in the time domain using least squares inverse filters and in the Fourier domain as the scale sum of the frequency components of the wavefields division.

Using the zero lag of the deconvolution imaging condition, we attenuate image artifacts due to wavefield multipathing through velocity anomalies. Also, we showed that it better handles amplitudes during imaging. The preceding statements are corroborated using two different synthetic data sets: one based on the convolutional model and the other based on a wave equation modeling program.

ACKNOWLEDGMENTS

I would like to thank Morgan Brown for his Fortran 90 programming assistance and Gabriel Alvarez for many discussions about deconvolution.

REFERENCES

Biondi, B., 2002, Prestack imaging of overturned and prismatic reflections by reverse time migration: SEP–**111**, 123–139.

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

Etgen, J., 1986, Prestack reverse time migration of shot profiles: SEP–**50**, 151–170.

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

Lee, D., Mason, I. M., and Jackson, G. M., 1991, Split-step fourier shot-record migration with deconvolution imaging:, *in* Geophysics Soc. of Expl. Geophys., 56, 1786–1793.

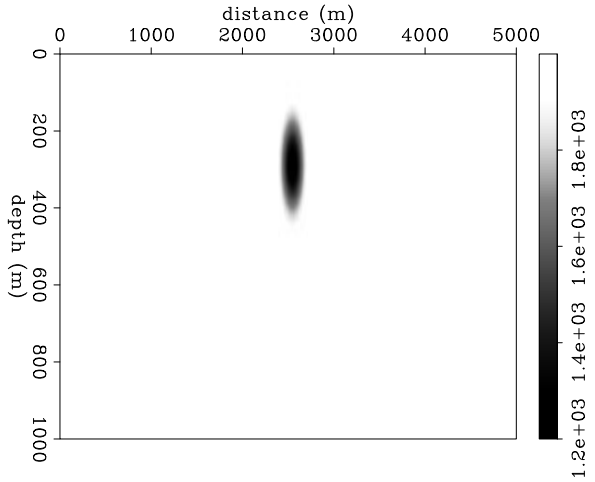


Figure 8: Velocity model with a Gaussian low velocity anomaly. alejandro1-velocity [ER]

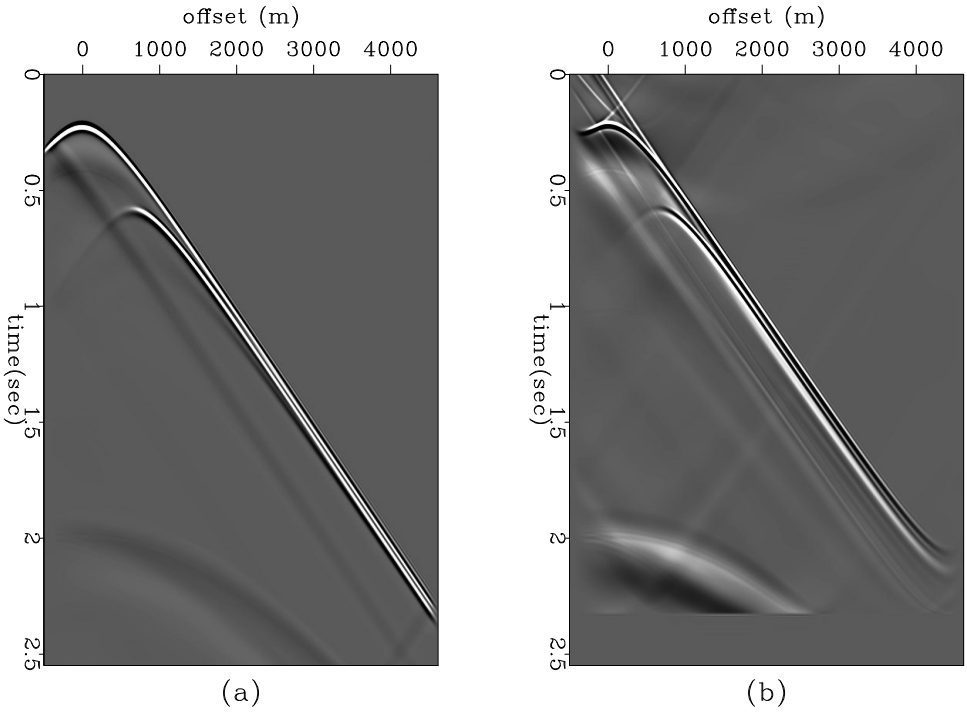


Figure 9: (a) Source wavefield. (b) Receiver wavefield at the reflector depth (700 m). alejandro1-shot-700 [ER]

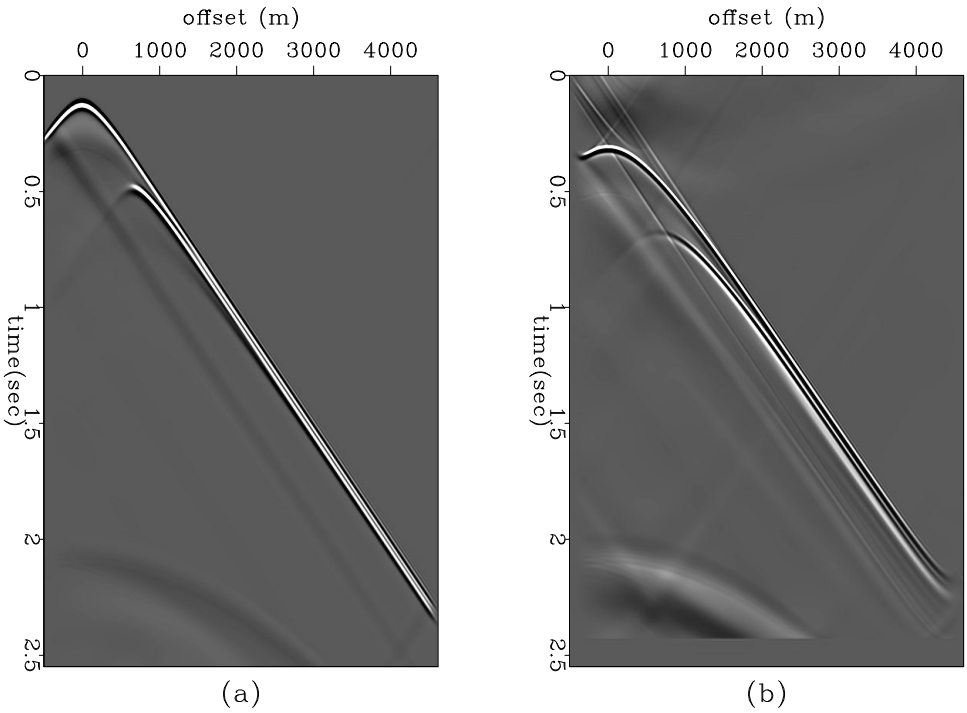


Figure 10: (a) Source wavefield. (b) Receiver wavefield at a shallower depth position (500 m). alejandro1-shot-500 [ER]

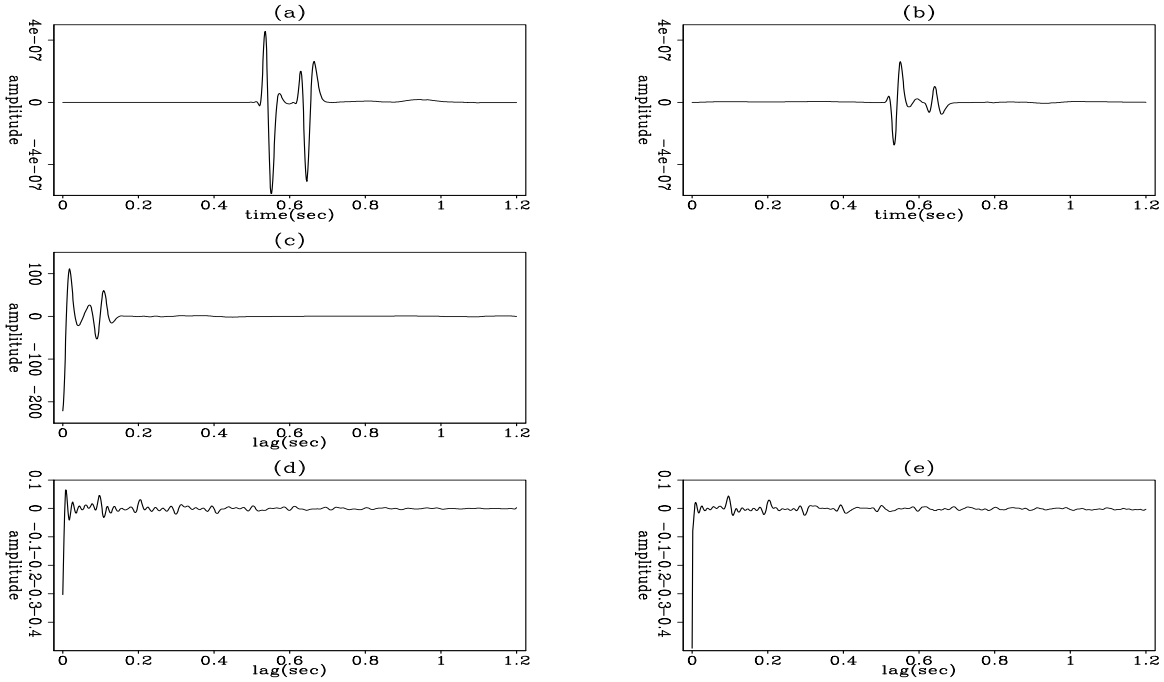


Figure 11: Source and receiver wavefields at the reflector depth at 1000 m offset. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-700 [ER]

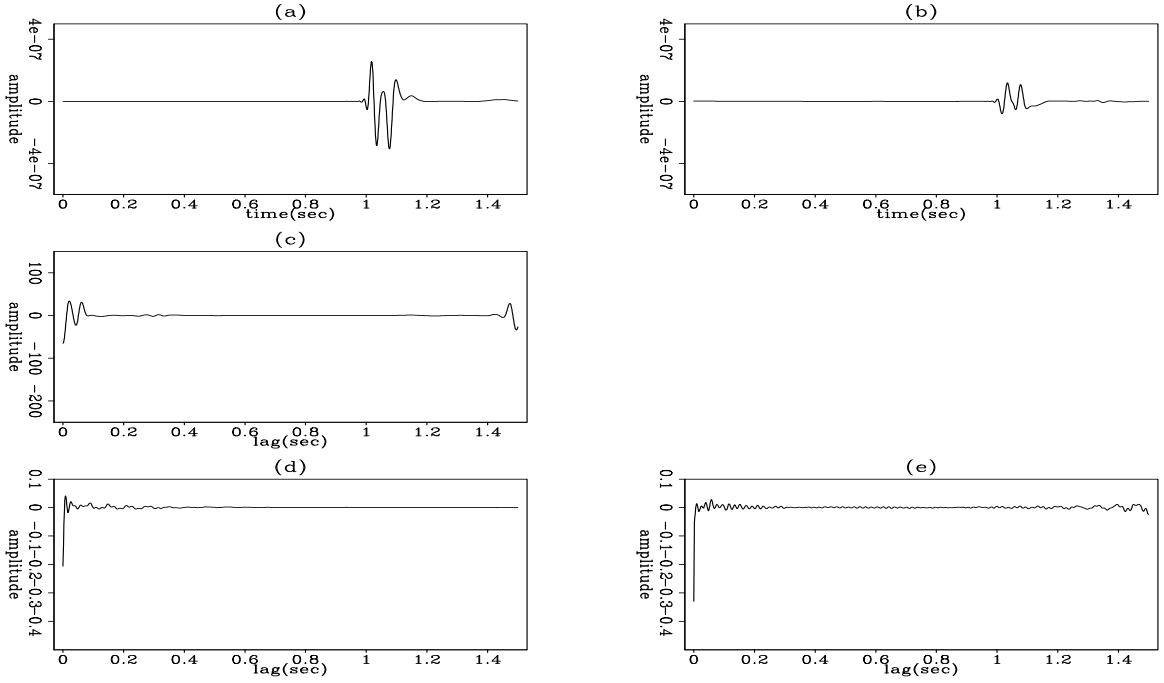


Figure 12: Source and receiver wavefields at the reflector depth at 2000 m offset. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-700-2000 [ER]

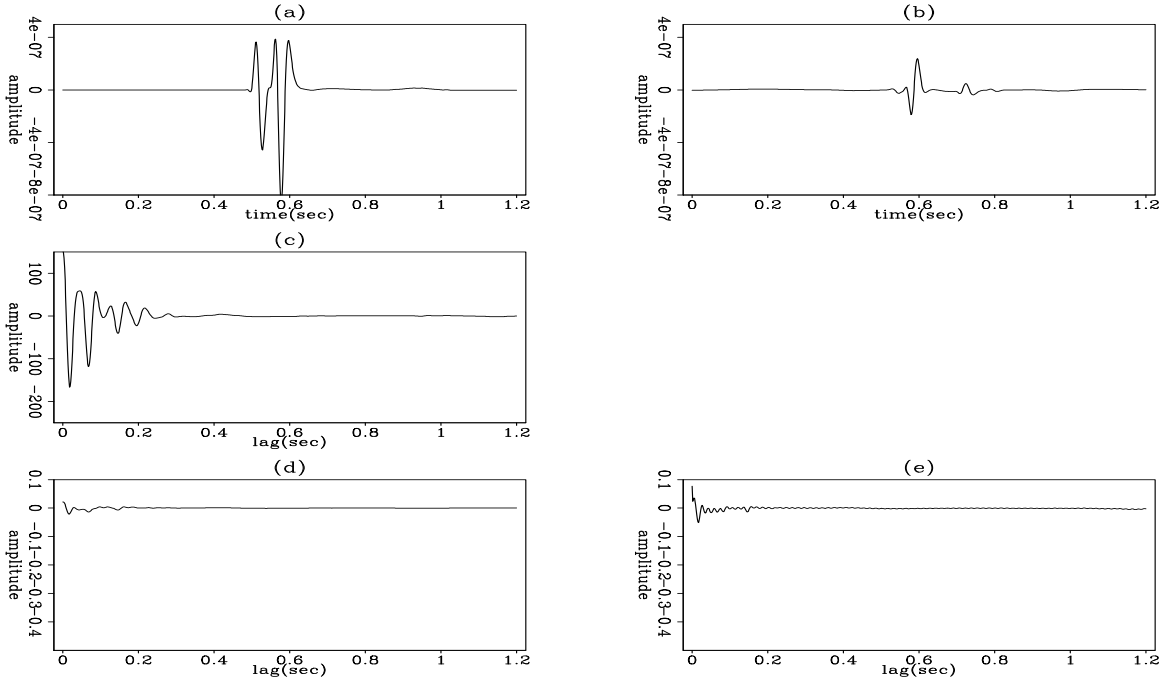


Figure 13: Source and receiver wavefields shallower than the reflector depth at 1000 m offset. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-500 [ER]

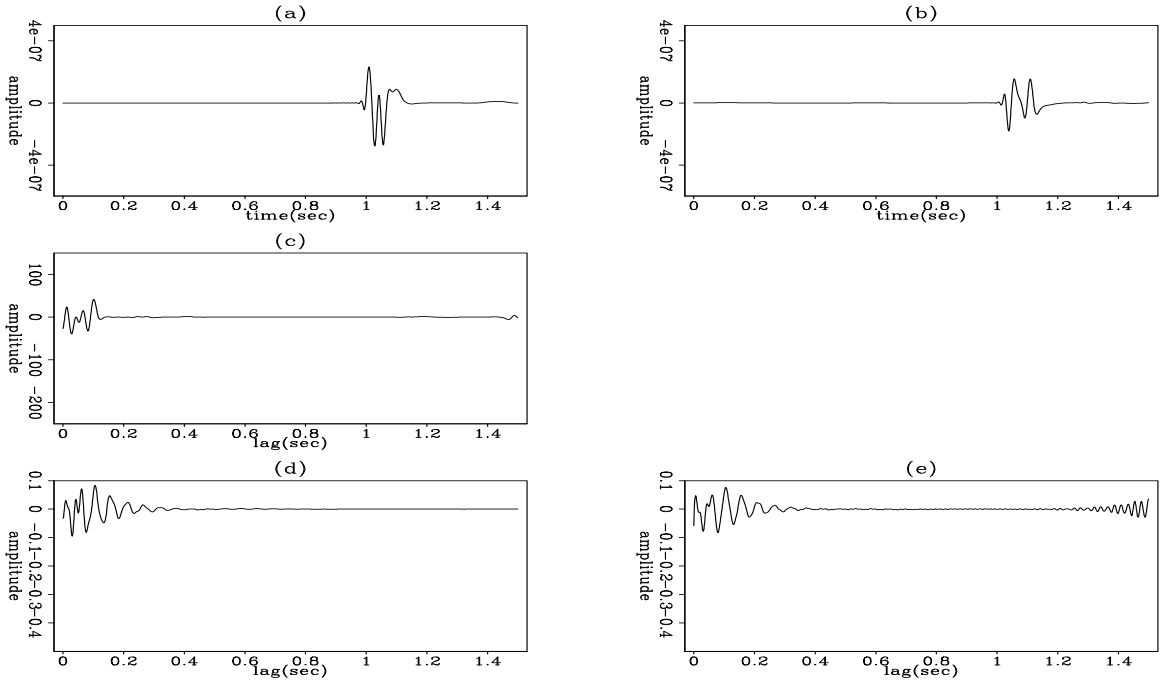


Figure 14: Source and receiver wavefields shallower than the reflector depth at 2000 m offset. (a) Source wavefield (b) Receiver wavefield (c) Cross-correlation (d) Deconvolution (b) by (a) in the time domain (e) Deconvolution (b) by (a) in the Fourier domain. alejandro1-500-2000 [ER]

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 97–109

Multicomponent data regularization

Daniel Rosales and Biondo Biondi [1]

ABSTRACT

Geometry regularization is a key process for obtaining reliable subsurface images with 3D seismic data. 3D regularization is, so far, a technique mainly used on *PP* land data. Multicomponent ocean bottom cable (OBC) technology simulates 3D land acquisition for multicomponent geophones at the ocean bottom. Reliable subsurface *PS* images, which provide amplitude information, have to go through a regularization process. Converted wave Azimuth Moveout (PSAMO) acts as a regularization operator in the formulation of the geometry regularization process in the least-squares sense.

INTRODUCTION

Multicomponent ocean bottom cable (OBC) technology reestablishes the use and importance of converted wave (*PS*) data, yet opens the door for a series of new and existing problems with *PS* data. Irregular acquisition geometries are a serious impediment for accurate subsurface imaging. Irregularly sampled data affects the image with amplitude artifacts and phase distortions. Irregular geometry problems are more evident in cases in which the amplitude information is one of the main goals of study. For *PS* data, this problem is crucial since most of the *PS* processing focuses on the estimation of rock properties from seismic amplitudes.

The application of inverse theory satisfactorily regularizes acquisition geometries of 3D prestack seismic data (Audebert, 2000; Chemingui, 1999; Duijndam et al., 2000; Rousseau et al., 2000; Albertin et al., 1999; Bloor et al., 1999; Nemeth et al., 1999; Duquet et al., 1998). For *PP* data, there are two distinct approaches to apply: 1) data regularization before migration and 2) irregular geometries correction during migration. Biondi and Vlad (2001) combine the advantages of the previous two approaches. Their methodology regularizes the data geometry before migration, filling in the acquisition gaps with a partial migration operator. The operator exploits the intrinsic correlation between prestack seismic traces. The partial migration operator used is Azimuth Moveout.

The recent development of a converted wave Azimuth Moveout (PSAMO) operator (Rosales and Biondi, 2001) that preserves amplitudes and is fast, enables the extension of Biondi and Vlad's (2001) methodology for converted waves data. Therefore, a complete and accurate

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

geometry regularization is now possible for OBC seismic data.

This paper extends already existing methodologies for *PP* regularization in order to handle *PS* data. Due to the asymmetry of ray trajectories in *PS* data, there are more elements to consider in order to solve for irregular geometry problems. Our method for *PS* data regularization uses a *PS* Azimuth Moveout operator (*PS*-AMO) (Rosales, 2002) in order to preserve the resolution of dipping events and correct for the lateral shift of the common conversion point.

Our methodology depends on the ratio between the *P* and the *S* velocities (γ). It also depends on the continuity of the events in the common midpoint gathers. These situations make our regularization an iterative procedure that stops where the difference between the previous and the actual γ sections is relatively small.

We will present a summary of Biondi and Vlad's (2001) methodology for solving the irregular geometry problem using a preconditioned-regularized least-squares scheme. We present and discuss how this method can be extended to handle *PS* data and implement this method on a portion of a real 3D OBC data set.

DATA REGULARIZATION

Regularized least-squares theory is the fundamental basis for solving the geometry regularization problem in this work. To preserve the resolution of dipping events in the final image, the regularization term includes a transformation by Azimuth Moveout (Biondi and Vlad, 2001). Additionally, Biondi and Vlad's method is computationally efficient because they apply the AMO operator in the Fourier domain and precondition the least-squares problem.

For this work, we use an AMO operator designed for converted waves (Rosales and Biondi, 2001). Regularization with this operator intends to: 1) preserve the resolution of the dipping events, 2) correct for the spatial lateral shift of the common conversion point, and 3) handle the amplitudes properly.

We present a general overview of the AMO regularization theory and discuss special considerations for converted waves regularization. We present an iterative methodology to regularize the *PS* data due to the dependency of the PSAMO operator on the ratio between the *P* and the *S* velocities.

AMO regularization overview

Partial stacking the data recorded with irregular geometries within offset and azimuth ranges yields uniformly sampled common offset/azimuth cubes. In order to enhance the signal and reduce the noise, the reflections should be coherent among the traces to be stacked. Normal Moveout (NMO) is a common method to create this coherency among the traces.

Let's define a simple linear model that links the recorded traces (at arbitrary midpoint locations) to the stacked volume (defined on a regular grid). Each data trace is the result of

interpolating the stacked traces and equal to the weighted sum of the neighboring stacked traces. In matrix notation, this transforms to:

$$\mathbf{d} = \mathbf{A}\mathbf{m}, \tag{1}$$

where $\mathbf{d}$ is the data space, $\mathbf{m}$ is the model space, and $\mathbf{A}$ is the linear interpolation operator. Stacking can be represented as the application of the adjoint operator $\mathbf{A}'$ to the data traces,

$$\mathbf{m} = \mathbf{A}'\mathbf{d}. \tag{2}$$

This simple operation does not yield satisfactory results for an uneven fold distribution. To compensate for this uneveness, it is common practice to normalize the stacked traces by the inverse of the fold ($\mathbf{W}_m$), thus:

$$\mathbf{m} = \mathbf{W}_m\mathbf{A}'\mathbf{d}. \tag{3}$$

Alternatively, it is possible to apply the general theory of inverse least-squares to the stacking normalization problem. The formal solution of the inverse least-squares problem takes the form:

$$\mathbf{m} = \left(\mathbf{A}'\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{d}. \tag{4}$$

Biondi and Vlad (2001) show that the fold normalization ($\mathbf{W}_m$) can be approximated as the inverse of $\mathbf{A}'\mathbf{A}$.

With the knowledge of model regularization in the least-squares inversion theory, it is possible to introduce smoothing along offset/azimuth in the model space. The simple least-squares problem becomes:

$$\begin{aligned} 0 &\approx \mathbf{d} - \mathbf{A}\mathbf{m} \\ 0 &\approx \epsilon_D \mathbf{D}'\mathbf{D}_\mathbf{h}\mathbf{m}, \end{aligned} \tag{5}$$

where the roughener operator $\mathbf{D}_h$ can be a leaky integration operator. However, the use of a leaky integration operator may yield the loss of resolution when geological dips are present. The substitution of the identity matrix in the lower diagonal of $\mathbf{D}_h$ with the AMO operator correctly transforms a common offset-azimuth cube into an equivalent cube with a different offset and azimuth. This transformation also preserves the geological dip.

The fold, which normalizes the data based on the traces distribution, is introduced by a diagonal scaling factor. The weights, for the regularized and preconditioned problem, are thus computed as:

$$\mathbf{W}_I{}^{-1} = \frac{\mathbf{diag}\left\{\left[\left(\mathbf{D_h}\mathbf{D_h}'\right)^{-1}\mathbf{A}'\mathbf{A}\left(\mathbf{D_h}'\mathbf{D_h}\right)^{-1} + \epsilon_D\mathbf{I}\right]\mathbf{p_{ref}}\right\}}{\mathbf{diag}(\mathbf{p_{ref}})}, \tag{6}$$

where $\mathbf{p_{ref}} = \mathbf{D_h}'\mathbf{D_h}\mathbf{m}$. This fold calculation can be simplified more as:

$$\mathbf{W_I}{}^{-1} = \frac{\mathbf{diag}\left\{\left[\left(\mathbf{D_h}\mathbf{D_h}'\right)^{-1}\mathbf{A}'\mathbf{A}\left(\mathbf{D_h}'\mathbf{D_h}\right)^{-1} + \epsilon_D\mathbf{I}\right]\mathbf{1}\right\}}{\mathbf{diag}(\mathbf{1})}. \tag{7}$$

PS regularization

We previously discussed that by formulating the irregular geometries problem in the least-squares sense it is possible to solve for gaps in the data using a regularization operator. The significant element of the previous section is the use of the AMO operator as the regularization term in the solution of the least-squares problem.

Recently, Rosales and Biondi (2001) developed and implemented an AMO operator for converted waves (PSAMO). This operator acts in the Fourier domain and also handles the amplitudes properly. Due to this new PSAMO operator, it is now possible to solve for the irregular geometries problem on converted wave data by following the same procedure as in the previous section.

Partial stacking requires the data to be coherent among the traces. NMO obtains this coherency well for *PP* data. However, for converted waves we know that the moveout is not a perfect hyperbola, even in constant velocity media.

On conventional *PP* processing, the AMO operator is velocity independent. However, for converted waves the PSAMO operator depends on the ratio between the *P* and the *S* velocities (γ). Therefore, we need *a priori* velocity estimation. This fact suggests that for different γ values we will have different regularization results.

Traditional *PS* processing intends to first sort the data in the common conversion point (CCP) domain. This process has always been dependent on the γ value; therefore, the *PS* processing community performs iterative processing (CCP binning, velocity analysis) until obtaining a satisfactory result.

The PSAMO operator that we use has the advantage of not demanding the data in the CCP domain. This operator is a cascade operation of converted wave dip moveout (Rosales et al., 2001) (PSDMO) and inverse PSDMO. The input for the PSDMO operator is in the CMP domain after NMO, since this operator performs the lateral shift correction.

After performing NMO on the *PS* data and the *PP* data, the γ value is (Huub Den Rooijen, 1991):

$$\gamma = \frac{v_p^2}{v_{eff}^2}, \tag{8}$$

where v_{eff} is the NMO velocity of the *PS* section.

In order to proceed with the *PS* data regularization, a process that depends on the γ value, we need to have the *PP* section regularized as well as the RMS velocity model. We proceed with the following algorithm:

1. Sort the data in the CMP domain.
2. Estimate velocity model on the *PS* section.
3. Estimate the γ section with equation (8).

4. If it is not the first iteration, compare the previous and the actual γ sections and:
   (a) if they are the same, finish the process.
   (b) if they are not, continue.
5. Apply NMO on the *PS* section.
6. Apply PSAMO regularization.
7. Apply inverse NMO.
8. Go back to step 2.

This is our main methodology to correctly regularize *PS* data.

RESULTS

We apply AMO regularization to a portion of a real OBC data set, the Alba field. 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 subsea (Newton and Flanagan, 1993).

Figure 1 illustrates the main problem. Observe the gaps in both the *PP*-CMP and the *PS*-CMP gathers. Our goal is to fill these gaps with energy from the surrounding traces and to preserve the physics of wave propagation.

A multicomponent OBC data set consists of both a *PP* and a *PS* section. Since the literature already presents extended work on *PP* regularization, we will present compact but complete regularization results for the *PP* portion. However, we will present more results and extended analysis for the *PS* section.

We use a portion of the entire 3D cube. This subsection consists of 17 crosslines with 719 cmps each. The *PP* section uses only the absolute value of the offset, for a total of 121 offsets. The *PS* section uses the full offset, however the maximum offset extension is reduced from 8000m. to 4000m. since the contribution of this far offset to this portion of the data is practically null.

PP regularization

Figure 2 presents the *PP* data for one crossline of the data set in study. Observe the holes in the data due to irregularities in the geometry acquisition.

Biondi and Vlad (2001) examined the differences among regularizing the data with normalization, regularization with the leaky integration operator and regularization with the AMO

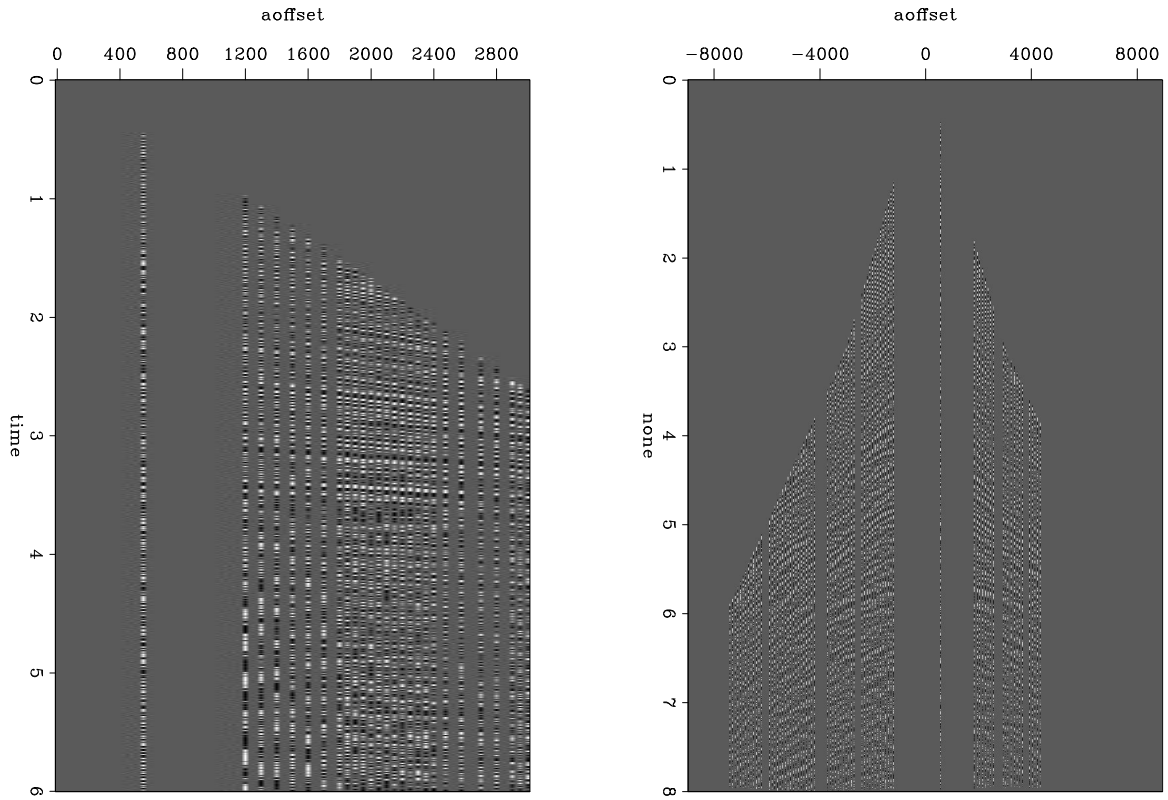


Figure 1: CMP gather for the *PP* (left) and the *PS* (right) components for one crossline of the 3D cube daniel2-cmps [ER]

operator. They conclude that the precondition of the regularized least-squares problem with the AMO operator yields more continuous results.

On this part of the problem, we only present the final interpolation results using normalization and AMO regularization. Figure 3 presents the fold maps calculated using both normalization (top) and AMO regularization (bottom). Note that even though the fold maps are similar, as expected, the fold distribution is smoother using AMO regularization. Also note that with AMO regularization, the fold reduces to the half. This fact affects the final solution of the least-squares problem.

Figure 4 compares the result of geometry regularization using normalization (top) and AMO regularization (bottom). Differences lie in the amplitudes and the borders.

PS regularization

Figure 5 exhibits the *PS* portion of the data set. Again, observe the holes in the data, as well as the presence of more offset.

The data is sorted into CMP gathers. We do not present the data in common conversion point (CCP) gathers because the input of the regularization program performs the binning, and

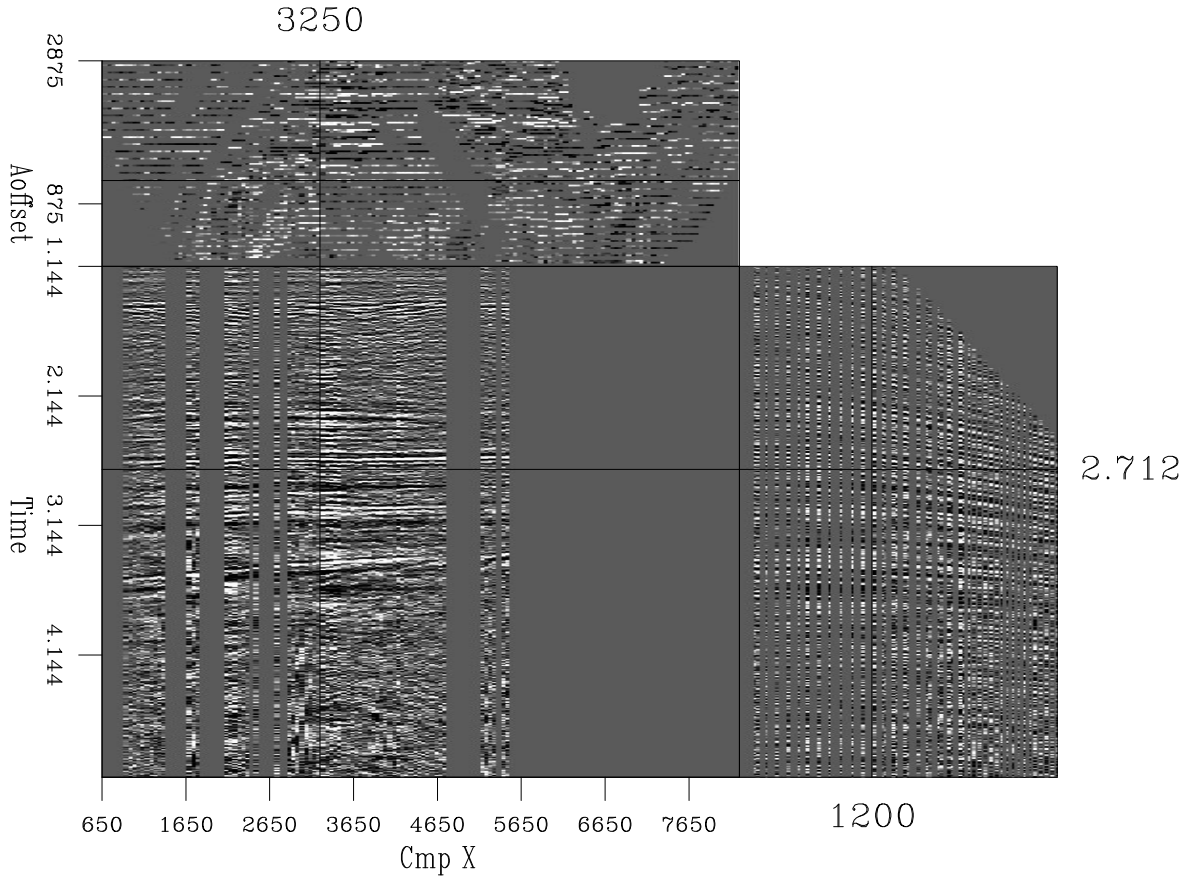


Figure 2: One *PP* crossline section of the data in study daniel2-data [ER]

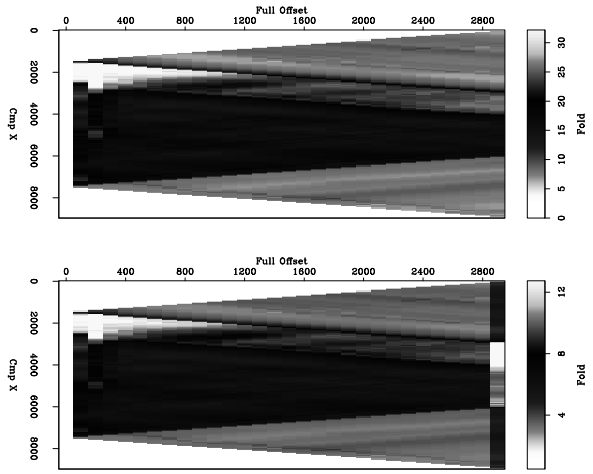


Figure 3: Fold, using normalization (top) and AMO regularization (bottom) daniel2-fold [CR]

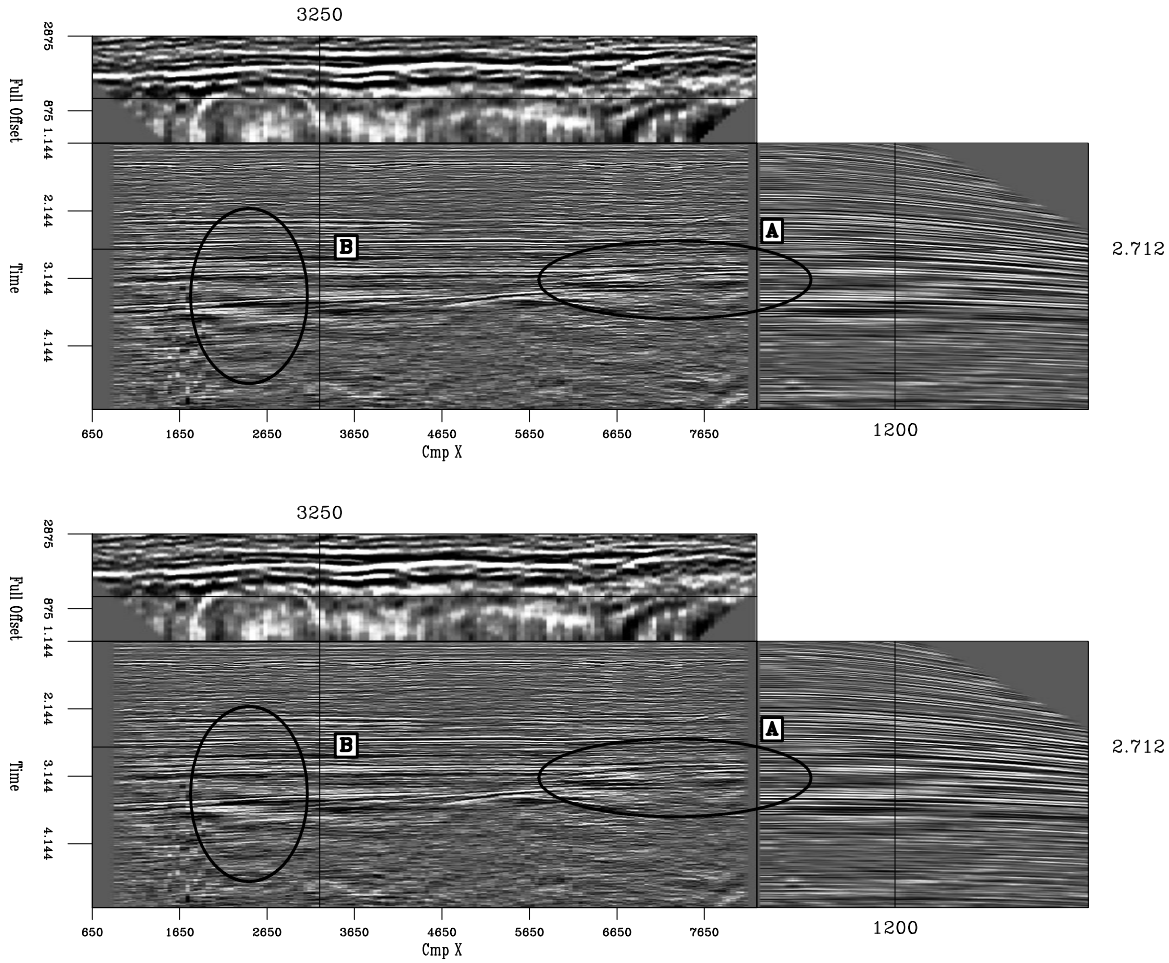


Figure 4: Data regularization results, using normalization (top) and the AMO operator (bottom). Note the main difference in the dipping events in the areas A and B daniel2-comp [CR,M]

the PSAMO operator performs the lateral shift correction from the CMP point to the CCP point based on the γ value.

We proceed with the methodology discussed in the previous section. We perform the PSAMO regularization process because it is the only one that corrects for the lateral shift displacement of the common conversion point. There are only two iterations so far.

Figure 6 presents two PSAMO regularization results for two iterations of our methodology. For each iteration note that the moveout of the events is not a perfect hyperbola. This characteristic corresponds to the nature of propagation of *PS* waves.

Observe the difference in the moveout of the events between the two iterations of our methodology (top and bottom parts of Figure 6). This is due to the different velocities on each iteration. However, both results satisfactorily fit the data.

Figure 7 present a zoom of our results and the original data. It is easier to observe that

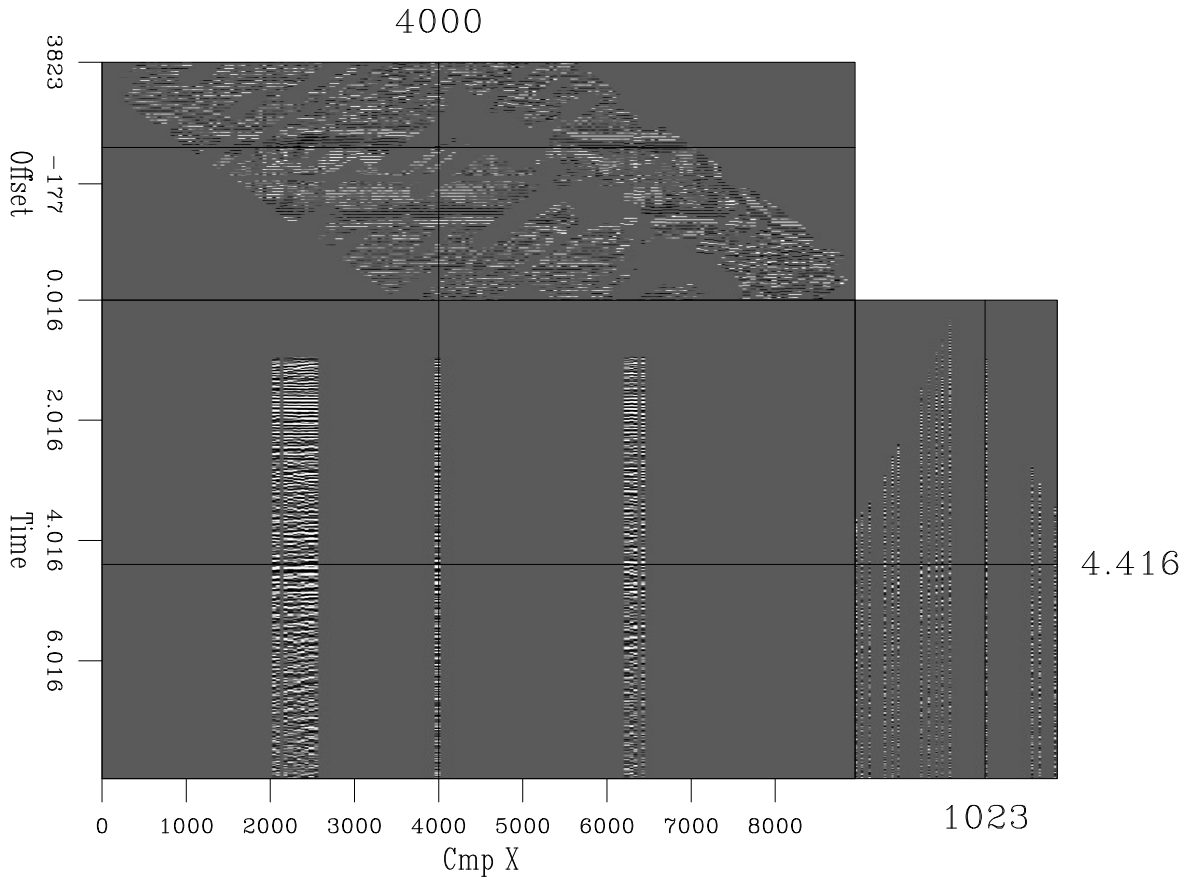


Figure 5: *PS* section for the same crossline on Figure 2 daniel2-data_ps [ER]

both results fit the data. However, the second iteration is more realistic since it better follows the information of the surrounding traces.

SUMMARY AND FUTURE WORK

We used least-squares inverse theory with the AMO operator as the regularization term. This method satisfactorily solved for interpolation of a 3D irregular data set.

We implemented a similar approach for regularizing the *PS* section of the OBC data set. For this problem, an iterative procedure is needed due to the dependence of the AMO operator on the γ value.

In order to obtain better results in the future, we recommend the use of a higher NMO approximation to obtain coherence among the traces to be stacked on the *PS* section. Additionally, formulating the γ estimation problem in a least-squares sense should allow a better constraint for its calculation, creating better *PS* regularized sections. This is an ongoing project.

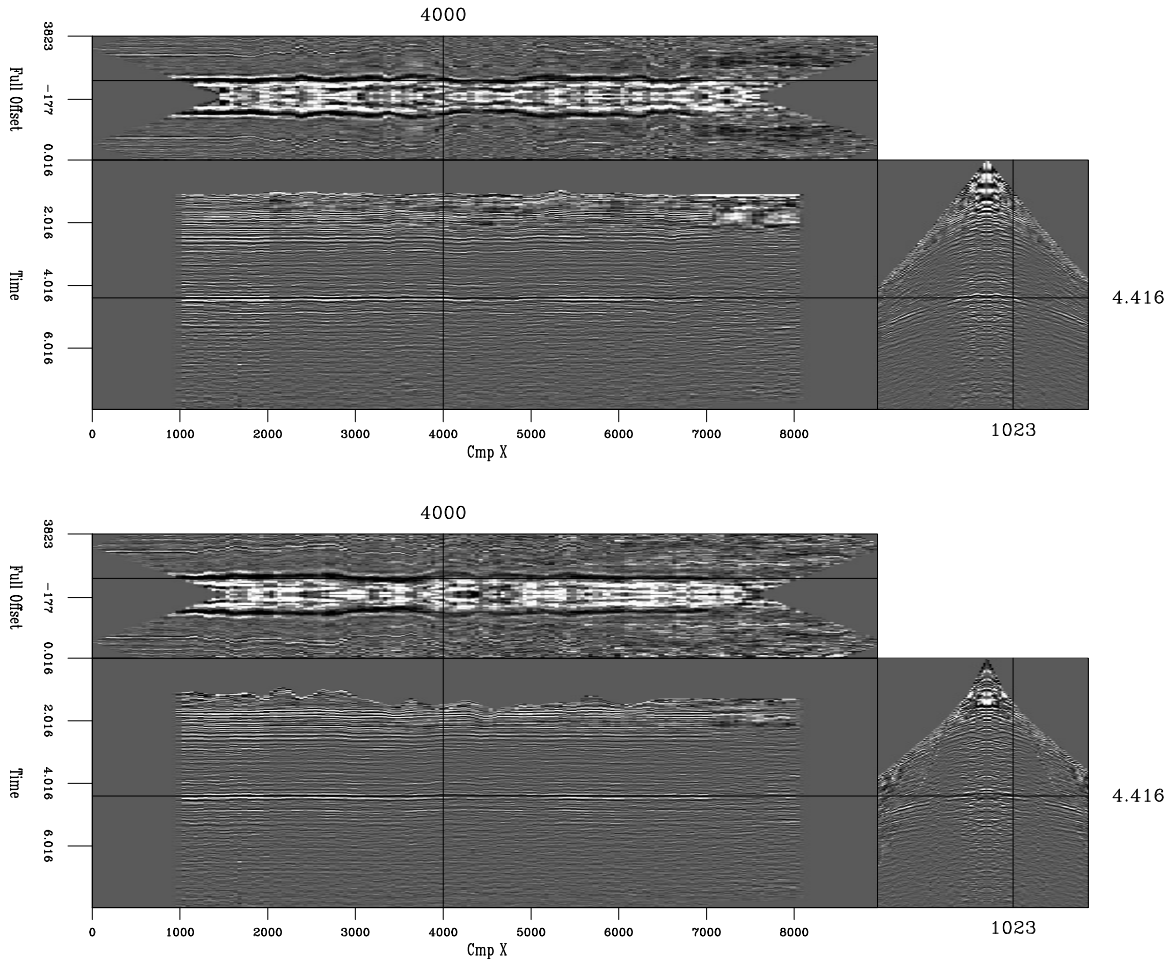


Figure 6: *PS* regularization results. First (top) and second (bottom) iteration of our methodology daniel2-comp_ps [CR,M]

ACKNOWLEDGMENTS

We want to thank ChevronTexaco for providing the data set.

REFERENCES

Albertin, U., Jaramillo, H., Yingst, D., Bloor, R., Chang, W., Beasley, C., and Mobley, E., 1999, Aspects of true amplitude migration: 69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1358–1361.

Audebert, F., 2000, Restored amplitudes AVA-gathers in the multi-angle domain: 62nd Mtg. Eur. Assoc. Expl Geophys., Abstracts, P–142.

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

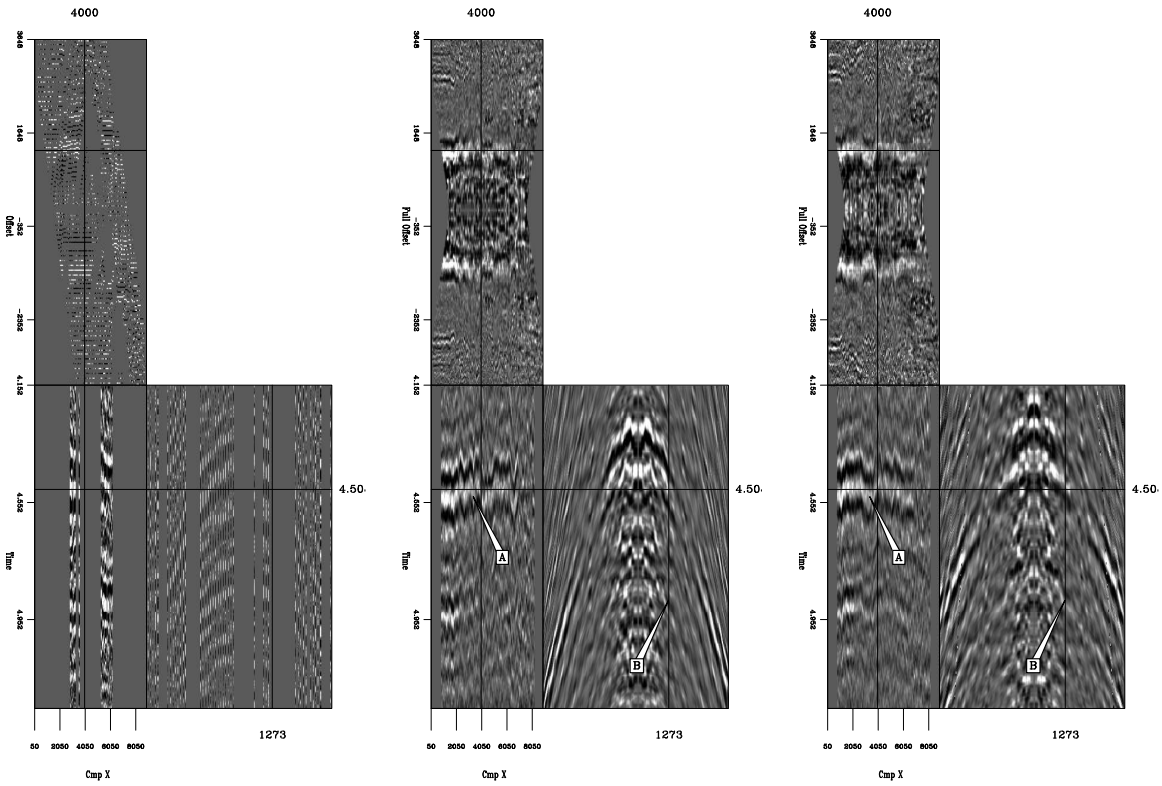


Figure 7: Zoom of Figure 6. From left to right, the original data, the first and the second iterations. Note how event A is more continuous in the second iteration, and event B presents a more realistic *PS* moveout (not a perfect hyperbola daniel2-detail [CR,M]

Bloor, R., Albertin, U., Jaramillo, H., and Yingst, D., 1999, Equalised prestack depth migration: 69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1358–1361.

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

Duijndam, A. J. W., Volker, A. W. F., and Zwartjes, P. M., 2000, Reconstruction as efficient alternative for least squares migration: 70th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, to be published.

Duquet, B., Marfurt, K., and Dellinger, J., 1998, Efficient estimates of subsurface illumination for Kirchhoff prestack depth migration: 68th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, 1541–1544.

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

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

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

Rosales, D., and Biondi, B., 2001, On asymmetric Stolt residual migration for imaging under salt edges: SEP–**110**, 71–74.

Rosales, D., Sava, P., and Biondi, B., 2001, Stolt residual migration for converted waves: SEP–**110**, 49–62.

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

Rousseau, V., Nicoletis, L., Svay-Lucas, J., and Rakotoarisoa, H., 2000, 3D true amplitude migration by regularisation in angle domain: 62nd Mtg. Eur. Assoc. Expl Geophys., Abstracts, B–13.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 109–117

Short Note

Equivalence of source-receiver migration and shot-profile migration

Biondo Biondi[1]

INTRODUCTION

At first glance, shot profile migration and source-receiver (survey-sinking) migration seem to be substantially different algorithms. The basic principles used by the two schemes are different. Shot profile migration is performed by *independently* propagating the source wavefield and the receiver wavefield. The image is obtained by cross-correlating (possibly normalized by the amplitude of the source wavefield) the two wavefields. Source-receiver migration is based on the concept of survey sinking, by which we recursively synthesize equivalent data sets at increasing depth. At each depth step imaging is performed by extracting the wavefield at zero time.

The issue of the relation between shot-profile migration and source-receiver migration has become more relevant since the recent introduction of methods for computing angle-domain common image gathers for source-receiver migration (Prucha et al., 1999; Sava et al., 2001). Rickett and Sava (2002) extended one of these methods [(Sava et al., 2001)] to downward-continuation shot-profile migration, and Biondi and Shan (2002) extended it to reverse-time shot-profile migration. Their extensions depend on the "equivalence" of the offset-domain common image gathers computed by shot-profile migration and source-receiver migration.

In this short note I demonstrate that the two migration methods produce exactly the same image cube; that is, the images are the same not only at zero subsurface offset, but also at non-zero subsurface offset. Wapenaar and Berkhout (1987) had already demonstrated the same result. Their focus, however, was on the stacked image, not on the whole image cube.

For the identity of the two methods to hold, the shot profile migration needs to satisfy three specific requirements: 1) the source function is an impulse at zero time and it has no spatial width, 2) the imaging condition is the cross-correlation of the source wavefield by the receiver wavefield, 3) the source and receiver wavefields are propagated by downward continuation. Another obvious assumption is that the same numerical algorithm is employed to downward continue the wavefields for both migration methods.

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

The demonstration of the equivalence of the two migration methods becomes fairly simple when we consider the migration of a single shot record by source-receiver downward continuation. In this case, the downward continuation of the sources is equivalent to the multiplication of the downward-continued receiver wavefield with many copies of the complex conjugate (time reversed) source wavefield appropriately shifted along the receiver axis.

THEORY

I first review the basic principles of shot-profile migration and source-receiver migration. Then I will show their equivalence.

Shot-profile migration

In shot-profile migration each record is migrated independently. The receiver wavefield P^g is downward continued starting from the recorded data. The source wavefield $P^s$is downward continued starting from an assumed source wavelet. In this case, we assume that the source is a delta function in the space domain and a constant as a function of frequency (impulse at time zero).

Each wavefield is propagated independently by convolution with the Single Square Root operator (SSR). At each depth level that is;

$$ {}^{\delta}P_z^s(\omega,x,y;\mathbf{s}) = \delta(x-x_s,y-y_s) \overset{x,y}{*} e^{-ik_z z}. \tag{1}$$

and

$$ P_z^g(\omega,x,y;\mathbf{s}) = P_{z=0}^g(\omega,x,y;\mathbf{s}) \overset{x,y}{*} e^{ik_z z}, \tag{2}$$

where x and y are defined in the image (model) space, and $\mathbf{s}=(x_s,y_s)$ is the location of the shot. The prefix δ in ${}^{\delta}P_z^s$ indicates that the source function is an impulse. Notice the negative sign in front of the exponential in equation (1). The negative sign is there because the source wavefield propagates downward, as opposed to propagate upward as the receiver wavefield does.

The image cube is formed by cross-correlating along the time axis the two wavefields shifted with respect to each other along the horizontal axes. In the frequency domain the cross-correlation is performed by multiplication with the complex conjugate, and it is evaluated at zero lag by summation over frequencies. The horizontal shift is the subsurface offset (x_h,y_h). The image cube is thus computed as:

$$ I_{\text{shot}}(z,x,y,x_h,y_h) = \sum_{x_s}\sum_{y_s}\sum_{\omega} P_z^g(\omega,x+x_h,y+y_h;\bar{\mathbf{s}})\overline{{}^{\delta}P_z^s(\omega,x-x_h,y-y_h;\bar{\mathbf{s}})} \tag{3}$$

Source-receiver migration

Source-receiver migration is based on the concept of survey sinking. After each depth propagation step, the propagated wavefield is equivalent to the data that would have been recorded

if all sources and receivers were placed at the new depth level. This task is accomplished by downward continuing all the source and receiver gathers at each depth step. Therefore, the basic downward continuation is performed by convolving with the Double Square Root (DSR) equation, as

$$P_z(\omega,\mathbf{g},\mathbf{s}) = P_{z=0}(\omega,\mathbf{g},\mathbf{s}) \overset{\mathbf{g}}{*} e^{ik_z z} \overset{\mathbf{s}}{*} e^{ik_z z}, \tag{4}$$

where the first convolution downward-continues the receiver wavefield, whereas the second convolution downward-continues the receiver wavefield. Notice the positive sign on both exponentials in equation (4).

At each depth level, the image is extracted from the downward-continued wavefield by evaluating the wavefield at zero time. The image-space coordinates and the source-receiver coordinates are linked by the well-known transformations

$$\begin{aligned} x_s &= x - x_h \qquad & x_g &= x + x_h \\ y_s &= y - y_h \qquad & y_g &= y + y_h. \end{aligned} \tag{5}$$

The image cube is then computed as

$$I_{\mathrm{s-g}}(z,x,y,x_h,y_h) = \sum_{\omega} P_z(\omega, x+x_h, y+y_h, x-x_h, y-y_h) \tag{6}$$

Equivalence of source-receiver migration and shot-profile migration

For the sake of simplicity, I demonstrate the equivalence by showing that the images obtained by migrating a single shot record are the same. The linearity of both migrations with respect to the input wavefield makes the extension to the full data set obvious.

A crucial observation for proving the equivalence of the two migration methods is that the downward continuation of the sources commutes with the downward continuation of the receivers. This property is obvious for vertically layered media where downward continuation can be performed in the wavenumber domain. However, it is also valid in presence of lateral velocity variations, because the wavefield is downward-continued along each direction by a convolution that is independent from the other direction. For example, the sources are downward continued by convolving each receiver gather with a convolutional operator that is non-stationary along the source axis, but is independent of the location of the receiver gather.

The wavefield at the surface for one single shot gather is given by the products of two functions: the first is independent of the source-coordinate $\mathbf{s}$ (the recorded data $P^g_{z=0}(\omega,\mathbf{g};\bar{\mathbf{s}})$), the second is independent of the receiver-coordinate $\mathbf{g}$ (a delta function at $\bar{\mathbf{s}}$).

The wavefield at depth obtained by survey sinking can thus be expressed as

$$\begin{aligned} & P_z(z,\mathbf{g},\mathbf{s}) \\ &= \left[P^g_{z=0}(\omega,\mathbf{g};\bar{\mathbf{s}})\,\delta(\mathbf{s}-\bar{\mathbf{s}})\right] \overset{\mathbf{g}}{*} e^{ik_z z} \overset{\mathbf{s}}{*} e^{ik_z z} \end{aligned}$$

$$= \left[P^g_{z=0}(\omega,\mathbf{g};\bar{\mathbf{s}}) \overset{\mathbf{g}}{*} e^{ik_z z}\right]\left[\delta(\mathbf{s}-\bar{\mathbf{s}}) \overset{\mathbf{s}}{*} e^{ik_z z}\right]$$
$$= \left[P^g_z(\omega,\mathbf{g};\bar{\mathbf{s}})\right]\left[\overline{{}^{\delta}P^s_z(\omega,x_s,y_s;\bar{\mathbf{s}})}\right]. \quad (7)$$

Imaging is performed by evaluating the downward-continued wavefield at the appropriate locations, as described by equation (6). If we apply this imaging condition to the wavefield in (7), we obtain

$$I_{\mathrm{s-g}}(z,x,y,x_h,y_h) = \sum_{\omega}\left[P^g_z(\omega,x+x_h,y+y_h;\bar{\mathbf{s}})\right]\left[\overline{{}^{\delta}P^s_z(\omega,x-x_h,y-y_h;\bar{\mathbf{s}})}\right], \quad (8)$$

that is exactly the same image cube as the image cube obtained by shot-profile migration [equation (3)].

TESTS ON SYNTHETIC DATA SET

To test the theoretical result reached in the previous section I migrated one shot record from a synthetic data modeled over a medium with strong lateral velocity variations that cause the image produced by migrating a single shot to have coherent artifacts. The test is useful both to demonstrate that in the source-receiver migration the downward continuation along the shot axis and along the receiver axis do indeed commute in presence of strong lateral variations. It is also interesting to confirm that the migration artifacts produced by source-receiver migration and shot profile migration are the same. The data set was kindly provided by Bill Symes of Rice University, and it has been used to study the artifacts produced by different kind of migrations (Stolk and Symes, 2002). The reflector geometry is a simple flat reflector, but a strong velocity anomaly above it creates severe multipathing that challenges different migration schemes.

The theoretical result is based on the assumption that the numerical algorithm used to propagate the wavefield is exactly the same. Unfortunately, SEP still lacks a downward-continuation shot-profile migration capable of handling severe velocity variations, though Brad Artman is close to succeeding in getting one up and running (2002). Therefore, for the moment I had to run an "imperfect" test and I compare the results of migrating synthetic data by source-receiver downward continuation and reverse-time shot-profile migration (Biondi and Shan, 2002).

Figure 1 shows the shot with source location at .5 kilometers used for the test. Figure 2 shows the zero offset (stack) image produced by both migrations methods. The panel on the left shows the image produced by shot-profile migration, and the panel on the right shows the image produced by source-receiver migration. The two images are similar, except for a small difference in frequency content caused by the fact that I did not enter a perfect impulsive source in the shot-profile migration to avoid dispersion. Not only the flat reflector is imaged similarly in the two images, but also the strong "ghost" reflectors caused by the triplication of the wavepath (Valenciano and Biondi, 2002), visible between the surface locations of 0 and 1 km, are almost identical.

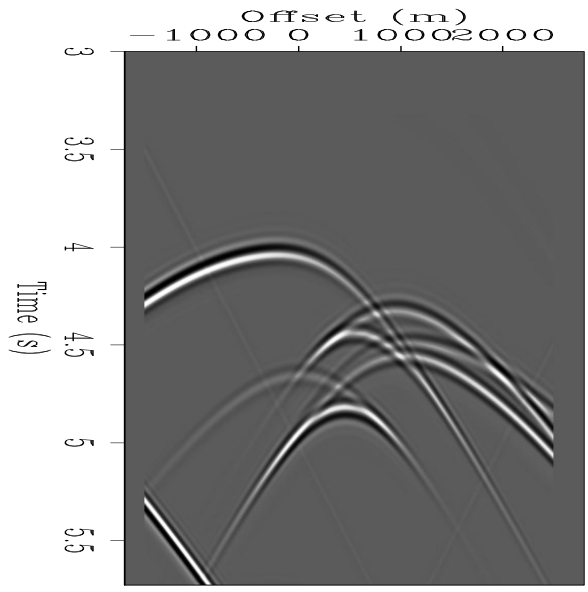


Figure 1: Shot profile used for the tests. biondo2-Shot-trip [CR]

Figure 3 shows the subsurface offset-domain common image gathers at the surface location of 300 meters: panel a) shot-profile migration, panel b) source-receiver migration). Again the images are similar for both the "true" reflector and the "ghost" reflectors. Figure 4 shows the angle-domain common image gathers obtained from the offset-domain gathers shown in Figure 3 after a slant stack transformation (Sava et al., 2001). Notice that the "true" reflector gets imaged at both positive and negative aperture angle because of the wavepath triplications. The "ghost" reflectors get imaged in the aperture-angle gap between the two branches of the true reflector.

Figure 5 demonstrates that the artifacts disappear if the whole data set (400 shots) is imaged. It shows the zero offset image [panel a)] and the angle-domain common image gather [panel b)] obtained by source-receiver migration when all the shot records are included in the data. I did not migrate all the shots by reverse time migration, because it would have taken considerable computer resources.

CONCLUSIONS

I have proven theoretically that source-receiver migration is exactly equivalent to downward-continuation shot-profile migration.

The results of the migration tests that I show strongly support this theoretical result, though they are not the ultimate proof, since I was limited to run a reverse-time shot-profile migration instead of a downward-continuation shot-profile migration.

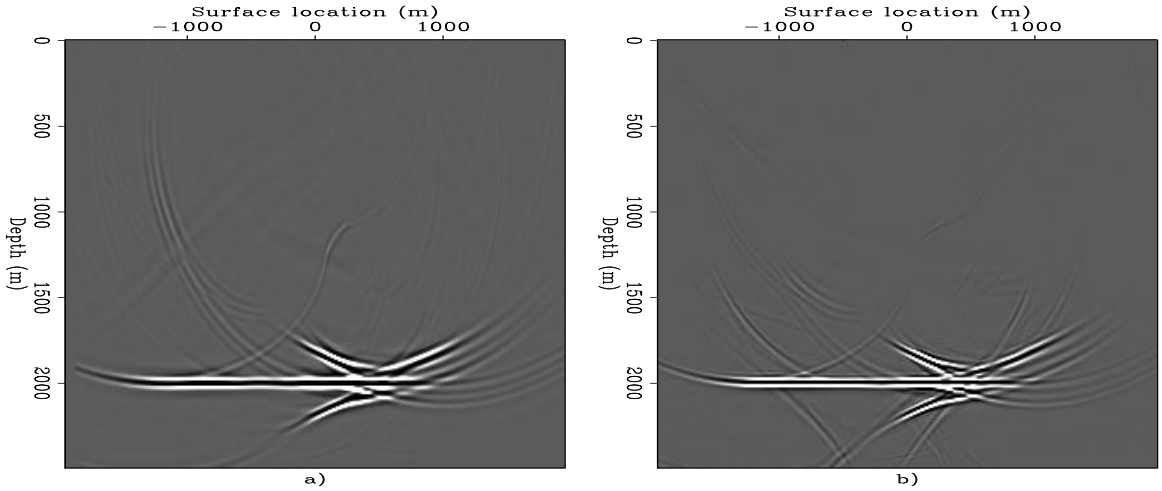


Figure 2: Zero-offset sections of the migrated cubes obtained using: a) shot-profile reverse-time migration, b) source-receiver downward-continuation migration. biondo2-Mig-trip-both [CR]

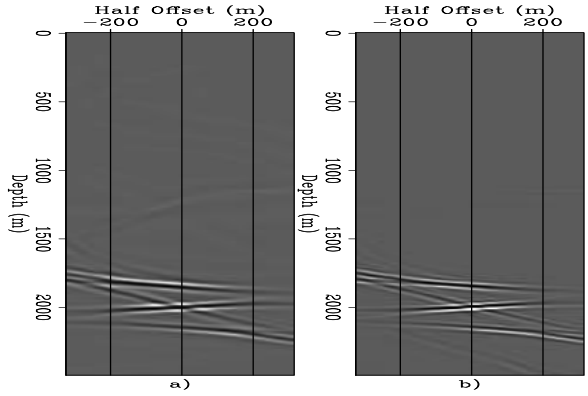


Figure 3: Offset-domain common image gathers obtained by slicing the migrated cubes at the surface location of 100 meters: a) shot-profile reverse-time image, b) source-receiver downward-continuation image. biondo2-Cig-trip-100-both [CR]

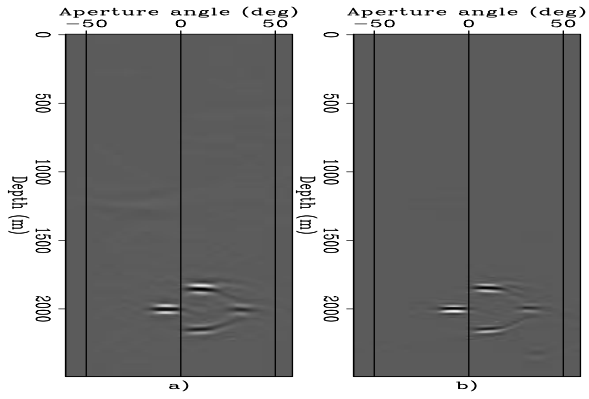


Figure 4: Angle-domain common image gathers obtained by slicing the migrated cubes at the surface location of 100 meters: a) shot-profile reverse-time image, b) source-receiver downward-continuation image. biondo2-Ang-Cig-trip-100-both [CR]

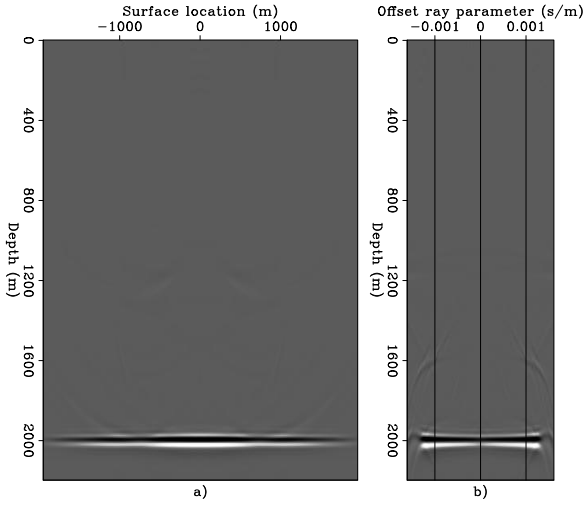


Figure 5: Stacked image (panel a)) and angle-domain common image gather at surface location of 100 meters (panel b)) obtained by source-receiver downward-continuation migration of the whole data set (400 shots). biondo2-Mig-Ang-Cig-trip-300-sr-full [CR]

ACKNOWLEDGMENTS

I would like to thank Bill Symes of Rice University for kindly providing the synthetic data set that was used for the tests shown in the paper.

REFERENCES

Artman, B., 2002, Migrating passive seismic data: SEP–**112**, 131–136.

Biondi, B., and Shan, G., 2002, Prestack imaging of overturned reflections by reverse time migration: 72nd Ann. Internat. Meeting, Soc. of Expl. Geophys., Expanded Abstracts, to be published.

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

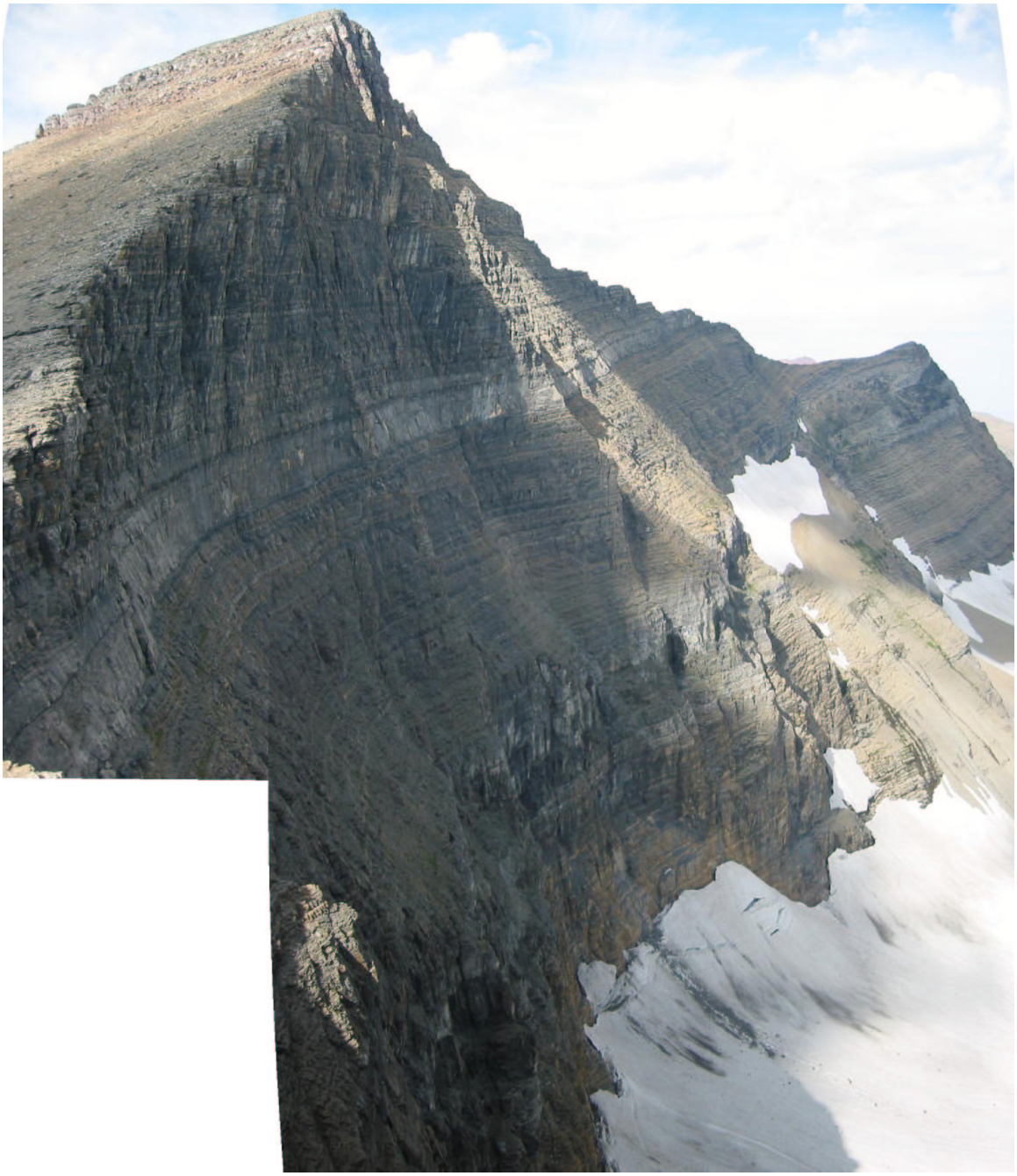
Rickett, J., and Sava, P., 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

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

Stolk, C., and Symes, W., 2002, Artifacts in Kirchhoff common image gathers: 72nd Ann. Internat. Meeting, Soc. of Expl. Geophys., Expanded Abstracts, to be published.

Valenciano, A. A., and Biondi, B., 2002, Deconvolution imaging condition for reverse-time migration: SEP–**112**, 83–96.

Wapenaar, C. P. A., and Berkhout, A. J., 1987, Full prestack versus shot record migration: 69th Ann. Internat. Meeting, Soc. of Expl. Geophys., Expanded Abstracts, Session:S15.7.



Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 117–123

Short Note

Damped imaging condition for reverse-time migration.

Alejandro A. Valenciano[1]

INTRODUCTION

Reverse-time migration of shot-profiles (Etgen, 1986) has been proposed (Biondi, 2002) as an alternative to downward continuation methods to perform imaging in complex subsurface environments (e.g., under complex and rugose salt bodies). In these situations, due to the poor illumination given by near-offset ray-paths, all the events present in the data (such as overturned reflections and prismatic reflections) are needed to generate interpretable images.

Not only the kinematic response of the migration is important, seismic data amplitudes also have the potential to provide information on reservoir properties. However, the most common implementation of shot-profile reverse-time migration uses the zero lag of the cross-correlation of the source and the receiver wavefields as imaging condition. This implementation has the advantage of being robust and honoring the kinematics of Claerbout's imaging principle (Claerbout, 1971) but does not honor the dynamics of the problem, resulting in the loss of amplitude accuracy.

I find that a damped imaging condition is more appropriate to obtain accurate amplitudes. I define the imaging condition as the zero lag of the cross-correlation of the source and the receiver wavefields divided by the sum of autocorrelation of the source wavefield and a constant damping factor. The division by the autocorrelation of the source wavefield acts as a normalization by the subsurface illumination.

The damping factor is useful because it avoids division by zero. Unfortunately, it introduces an error in the image amplitudes. I used a mask function that is inversely proportional to subsurface illumination to avoid the use of the damping when it is not needed (space and time variable damping).

Using a shot from a 3D marine seismic dataset acquired in a complex area, I compare three different imaging conditions: cross-correlation, division with constant damping, and division with variable damping. I find that the variable damping imaging condition preserves the amplitudes in areas with good subsurface illumination. In areas with poor subsurface illumination, it does the same job as the constant damping imaging condition.

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

DAMPING IN CONVENTIONAL IMAGING CONDITION

Conventional shot-profile migration determines the reflection strength at each subsurface point by taking into account only the source and receiver wavefield at that location. Jacobs (1982) compares two different imaging conditions

$$\mathbf{r} = \sum_{\mathbf{shot}} \sum_{\mathbf{t}} \mathbf{ud}, \tag{1}$$

and

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

The first is the most commonly used in 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 ε^2), but has the disadvantage of relying on a damping factor that cannot be automatically estimated (Claerbout, 1992).

MASKING THE DAMPING FACTOR

The damping factor is useful because it avoids instability in noisy-signal division but problematic because it biases the image amplitudes. I propose to add a mask function inversely proportional to the subsurface illumination at each point (Rickett, 2001).

$$\mathbf{w} \quad \alpha \quad \frac{\mathbf{1}}{\mathbf{d^2}}. \tag{3}$$

When $\mathbf{d^2}$ has enough energy to contribute to the image, the damping factor ε is set to zero. When factor $\mathbf{d^2}$ is small, the damping factor is kept to avoid zero division. Thus, the imaging condition can be set as

$$\mathbf{r} = \sum_{\mathbf{shot}} \sum_{\mathbf{t}} \frac{\mathbf{ud}}{\mathbf{d^2} + \mathbf{w}\varepsilon^2}, \tag{4}$$

where the damping is now variable in space and time.

RESULTS

A 2D shot from a seismic 3D marine dataset acquired in a complex area was used to test the preceding idea. Figure 1 shows the source wavefield and Figure 2 the receiver wavefield.

I calculated the reflection strength using the three imaging conditions stated in equations (1), (2) and (4). The results are shown in Figures 3-5. We can see that the damped imaging conditions from equation (2) and equation (4) give a more balanced section. Some artifacts, like the shot artifact present in Figure 3, were eliminated in Figure 4 and Figure 5. In general, the continuity of the events was enhanced.

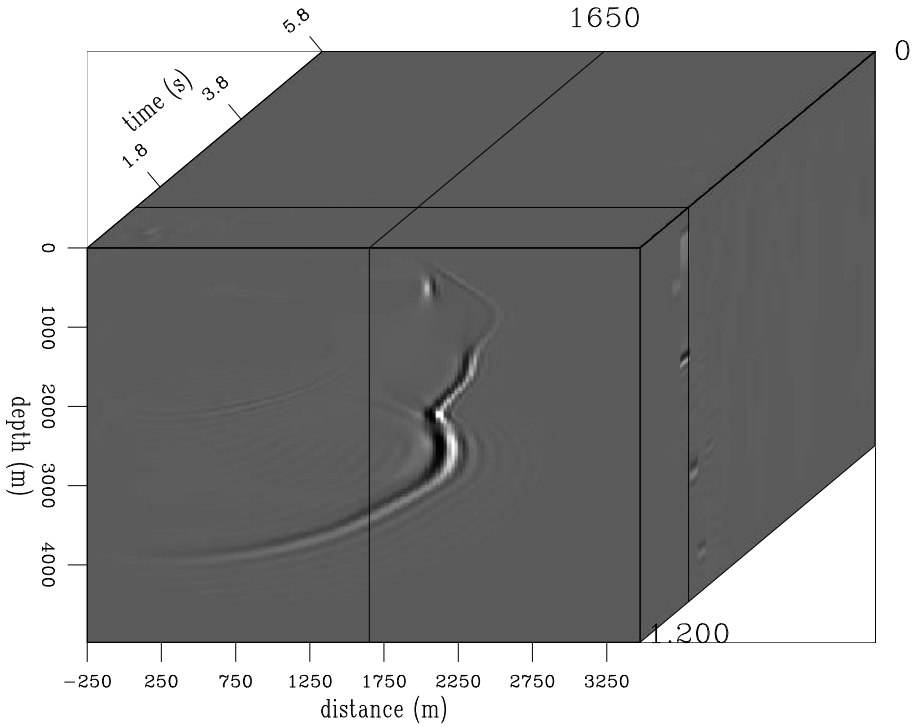


Figure 1: One shot source wavefield. alejandro2-D [ER]

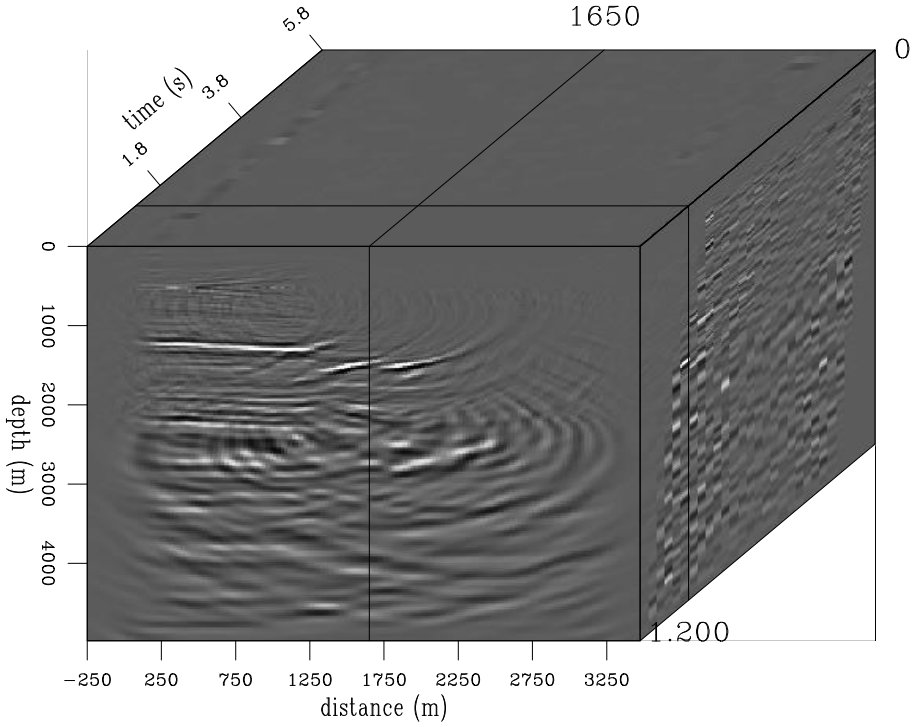


Figure 2: One shot receiver wavefield. alejandro2-U [ER]

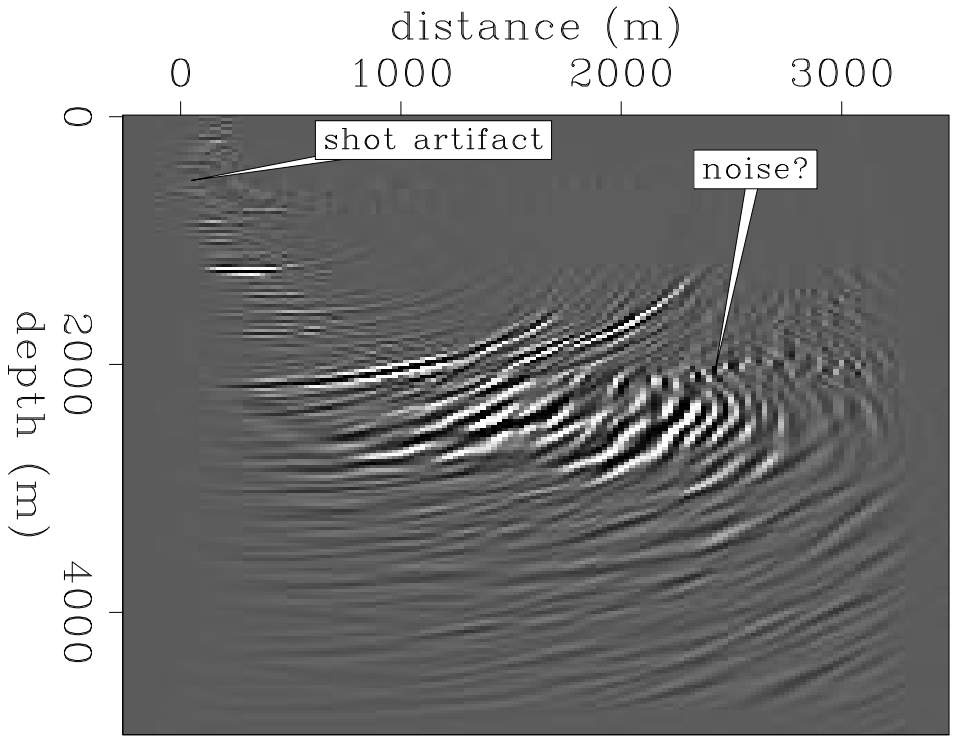


Figure 3: One shot stack using zero lag of the cross-correlation imaging condition as stated in equation (1). alejandro2-Image [ER]

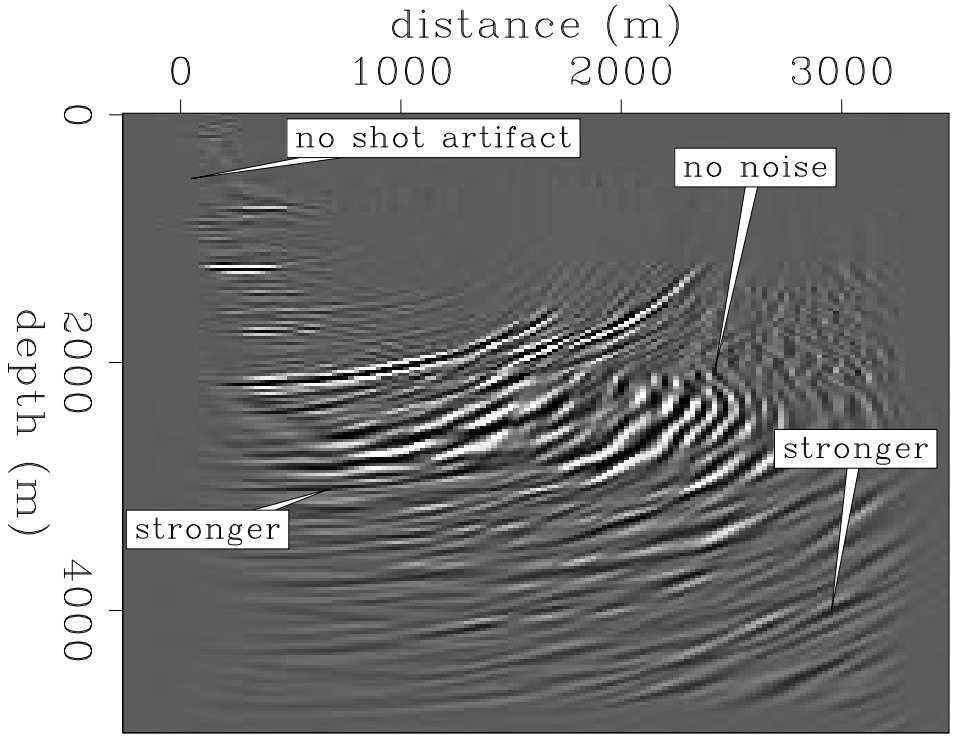


Figure 4: One shot stack using constant damped imaging condition as stated in equation (2). alejandro2-Image_damp [ER]

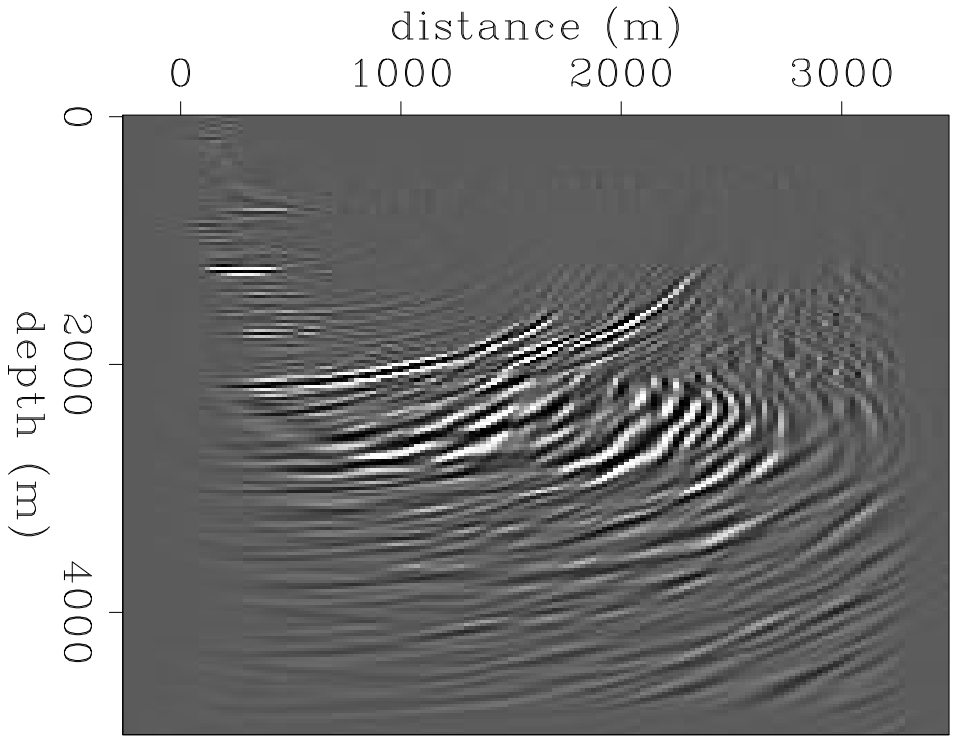


Figure 5: One shot stack using variable damped imaging condition as stated in equation (4). alejandro2-Image_damp1 [ER]

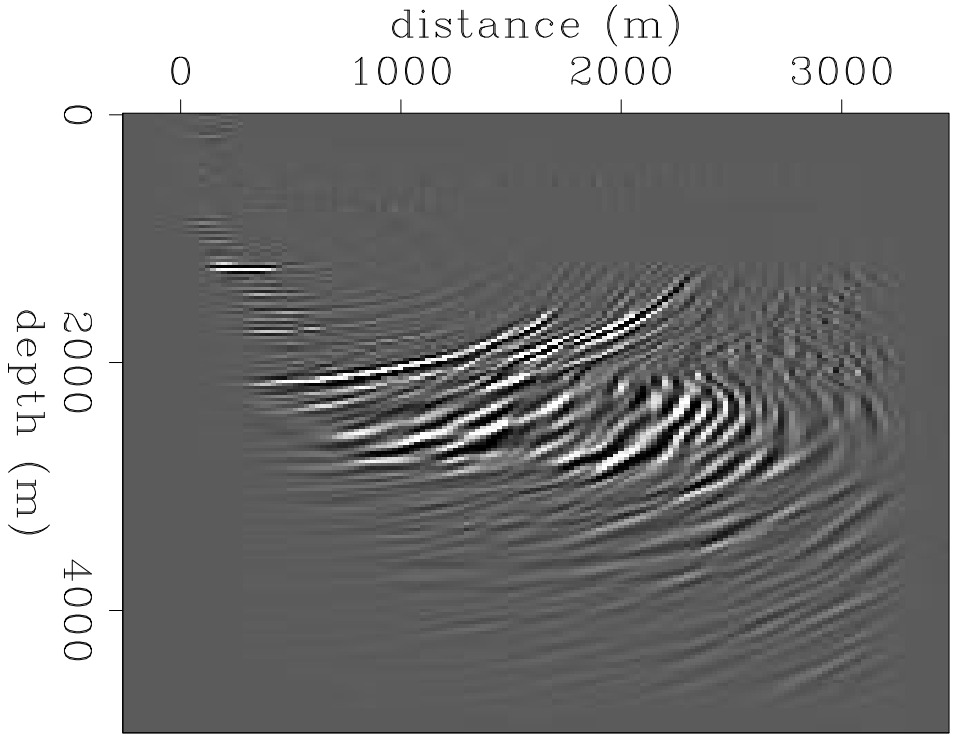


Figure 6: Difference between the constant damping (Figure 4) and the variable damping (Figure 5). alejandro2-diff_damp [ER]

In the two damped images compared in Figure 6, the difference between the images is greatest in the center where the damping was not needed and least in the corners where it is really needed.

A final test for the variable damping imaging condition should be the stack of more shots to form the final image. But issues like the data driven selection of the damping factor still need to be addressed to make it applicable to a full seismic dataset.

CONCLUSIONS

I showed, using a shot from a seismic 3D marine dataset acquired in a complex area, that including a damping factor in the imaging condition can improve the amplitude accuracy of the conventional shot-profile reverse-time migration imaging condition.

The damping factor can be related to reflector illumination, adding the damping factor where it is really needed. Thus a variable damping approach should better preserve the amplitudes than a constant damping approach.

ACKNOWLEDGMENTS

I would like to thank Guojian Shan for generating the data I used in this paper.

REFERENCES

Biondi, B., 2002, Prestack imaging of overturned and prismatic reflections by reverse time migration: SEP–**111**, 123–139.

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

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

Etgen, J., 1986, Prestack reverse time migration of shot profiles: SEP–**50**, 151–170.

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

Rickett, J., 2001, Spectral factorization of wavefields and wave operators: Ph.D. thesis, Stanford University.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 123–131

Short Note

Multicomponent Stolt residual migration: a real data example

Daniel Rosales [1]

INTRODUCTION

Residual migration is the process of updating an image without re-running the entire migration process. Its main application is in refining of velocity models. Rosales et al. (2001) first introduced Stolt residual migration for converted waves (*PS*).

Stolt migration is a constant velocity process. Although Stolt residual migration is also based on the assumption of constant velocity, Sava (2000) shows that Stolt residual migration is valid in variant velocity media, since the residual migration operator depends only on the ratio between the migration velocity and the real velocity.

For converted waves, Stolt residual migration basically depends on the combination of three out of four parameters: 1) The ratio between the migration and the real *P*-velocity (ρ_p), 2) the ratio between the migration and the real *S*-velocity (ρ_s), 3) the ratio between the *P* and *S* migration velocities (γ_0) and 4) the ratio between the real *P* and *S* velocities (γ).

Understanding how these parameters interact and affect *PS* residual migration in a non-constant velocity medium is of crucial importance for future velocity analysis studies. I present residual migration results on a 2D line extracted from the 3D OBC data set of the Alba field after geometry regularization (Rosales and Biondi, 2002).

THEORY REVIEW

Rosales et al. (2001) describe three possible ways to perform residual migration. The most precise method is an exact derivation which involves the combination of the ρ_p, ρ_s and γ_0 parameters. This method attempts to simultaneously correct the effect of two inaccurate velocity fields.

Assuming that the initial migration was done with the velocities v_{0p} and v_{0s}, and that the

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

correct velocities are v_{mp} and v_{ms}, we can then write

$$\begin{cases} k_{z_0} = \frac{1}{2}\left(\sqrt{\frac{\omega^2}{v_{0p}^2} - k_s^2} + \sqrt{\frac{\omega^2}{v_{0s}^2} - k_g^2}\right) \\ k_{z_m} = \frac{1}{2}\left(\sqrt{\frac{\omega^2}{v_{mp}^2} - k_s^2} + \sqrt{\frac{\omega^2}{v_{ms}^2} - k_g^2}\right). \end{cases} \tag{1}$$

Solving for ω^2 in the first equation of (1) and substituting it in the second equation of (1), we obtain the expression for prestack Stolt depth residual migration for converted waves:

$$k_{z_m} = \frac{1}{2}\sqrt{\rho_p^2 {\kappa_0}^2 - k_s^2} + \frac{1}{2}\sqrt{\rho_s^2 {\gamma_0}^2 {\kappa_0}^2 - k_g^2}, \tag{2}$$

where ${\kappa_0}^2$ is the transformation kernel defined as:

$${\kappa_0}^2 = \frac{4({\gamma_0}^2+1){k_{z_0}}^2 + ({\gamma_0}^2-1)({k_g}^2 - {k_s}^2) - 4k_{z_0}\sqrt{(1-{\gamma_0}^2)({\gamma_0}^2{k_s}^2 - {k_g}^2) + 4{\gamma_0}^2{k_{z_0}}^2}}{({\gamma_0}^2-1)^2},$$

and $\rho_p = \frac{v_{0p}}{v_{mp}}$, $\rho_s = \frac{v_{0s}}{v_{ms}}$, and $\gamma_0 = \frac{v_0 p}{v_0 s}$.

This formulation depends only on velocity ratios. This fact implies that it is a valid formulation for non-constant velocity media, as suggested by Sava (2000).

Methodology

Residual migration is a very useful tool for velocity analysis. Even though it is based on a constant velocity media, the residual migration depends only on velocity ratios. Therefore, it is approximately correct and applicable for variable velocity models, at least with depth migration.

Since I handle Stolt residual migration for converted waves using three parameters, it is important to simplify the problem. The more parameters we have to search for, the more complicated the problem is. Using the *PP* data in order to update the *P* velocity model the best possible way, will lead to only a two-parameter estimation (ρ_s and γ).

REAL DATA RESULTS

In order to implement and test Stolt residual migration for converted waves, I use a 2D line from a 3D cube of the OBC data set of the Alba oil field. The 2D line is extracted from the 3D portion after geometry regularization (Rosales and Biondi, 2002).

The velocity models were provided by Chevron. I start with these velocity models. I first handle the *PP* section, perform Stolt residual migration on this section only and update the velocity model with a selection of combined values of ρ_p. After the *P* velocity model is

updated, I perform the *PS* migration and residual migration; however, I do not scan over ρ_p values. I scan only over ρ_s and γ values.

The velocities vary both horizontally and vertically. Therefore, there will not be a single set of parameters to correct the entire section. The γ_0 value will be choosen at Alba's depth at the center of the section.

PP results

Figure 1 shows the migration result. Figure 2 shows three angle gathers. Observe the residual moveout in these angle gathers. This indicates errors in the velocity model.

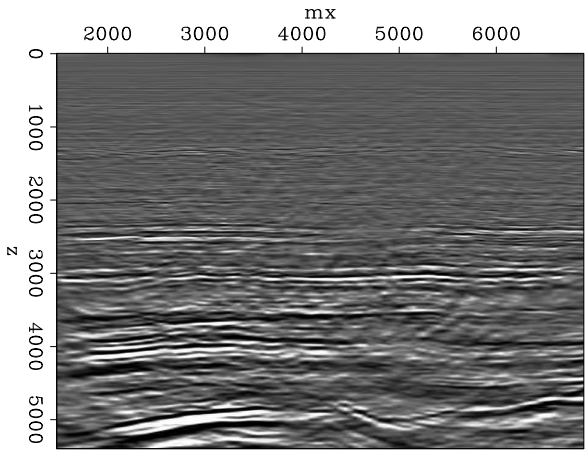


Figure 1: *PP* migration result with the original *P*-velocity model daniel1-ppmiga [CR]

I perform Stolt residual migration (Sava, 1999) over this migration result. There is not a single value of ρ_p that fixes the whole section. Figure 3 shows the same three angle domain common image gathers, as in Figure 2. Moreover, each angle domain common image gather corresponds to a different value of ρ_p, from left to right, $\rho_p = 0.98$ $\rho_p = 0.99$ and $\rho_p = 1.02$. Note how different events correct better for different ρ_p values.

By updating the velocity model with different ρ_p values at different depths and lateral positions, I perform a new migration. Figure 4 shows this final migration result. Figure 5 shows again the same different angle domain common image gathers for the migration result of Figure 4.

PS results

Figure 6 shows the *PS* migration result, using the updated *P* velocity model and the given *S* velocity model. Figure 7 shows three angle domain common image gathers for this migration result.

Figure 8 presents the residual migration results keeping the ρ_p value as 1, i.e, no changes in the *P*-velocity field, and a ρ_s value of 0.98. Figure 9 shows the angle domain common

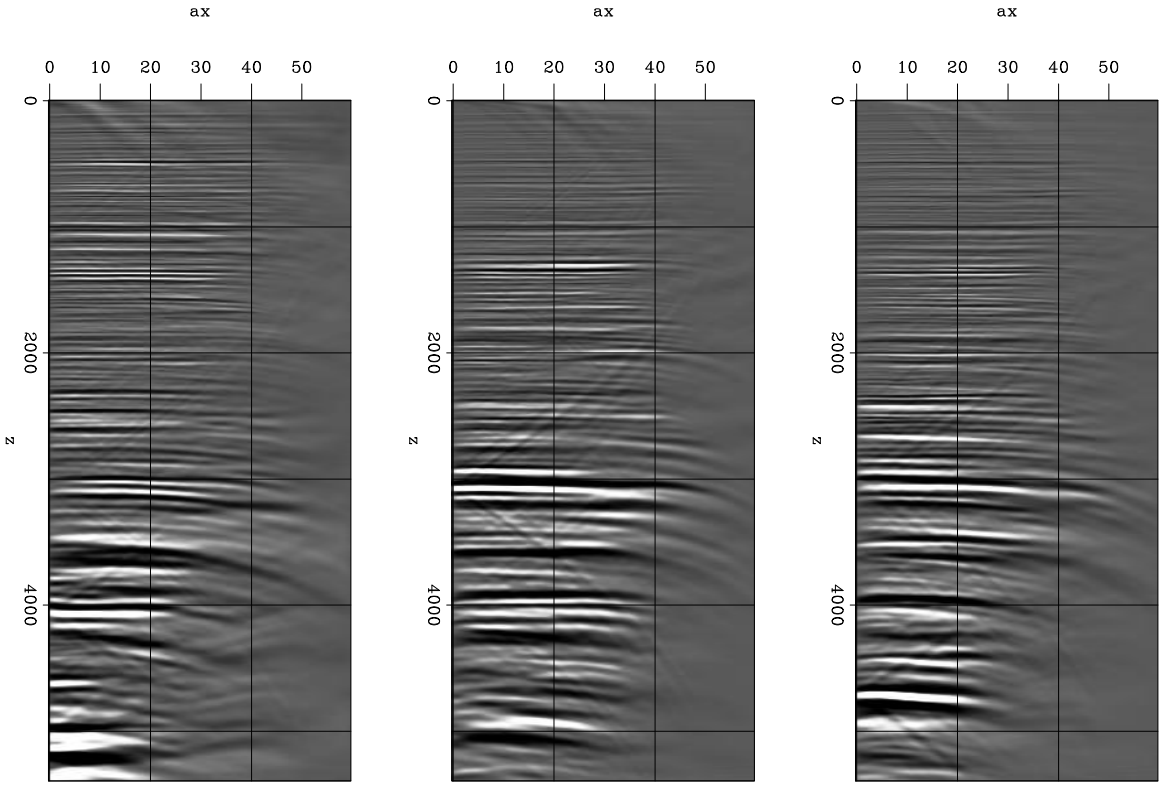


Figure 2: Three angle domain common image gathers for the migration result of Figure 1. They corresponds, from left to right, to a lateral position of 2500m, 4500m and 6500m. Observe the residual curvature. daniel1-ppmigb [CR,M]

image gathers for the residual migration process. These angle gathers are at the same position as in Figure 7, however, they correspond to ρ_s values of $\rho_s = 0.98$ $\rho_s = 0.99$ and $\rho_s = 1.02$.

The γ_0 value used for the PS residual migration is constant for the whole section, with a value of $\gamma_0 = 1.728709$. The γ_0 value was chosen at a depth of 1900m, which is the estimated depth of the Alba reservoir, and at the center of the section.

SUMMARY

I present the application of Stolt residual migration for converted waves on a real data set. Many challenges are involved in this application. The most important is the sparsity of the data. This problem is well addressed by Rosales and Biondi (2002).

In order to handle a 6D cube of information for a 2D prestack line after residual migration for converted waves, I freeze the ρ_p parameter after applying residual migration only on the *PP* section. A more robust technique to deal with all the parameters simultaneously is in progress.

Picking the parameter values is also a troublesome problem, Clapp (2002) presents some suggestions on how to deal with this problem on *PP* data only. A simultaneous picking might

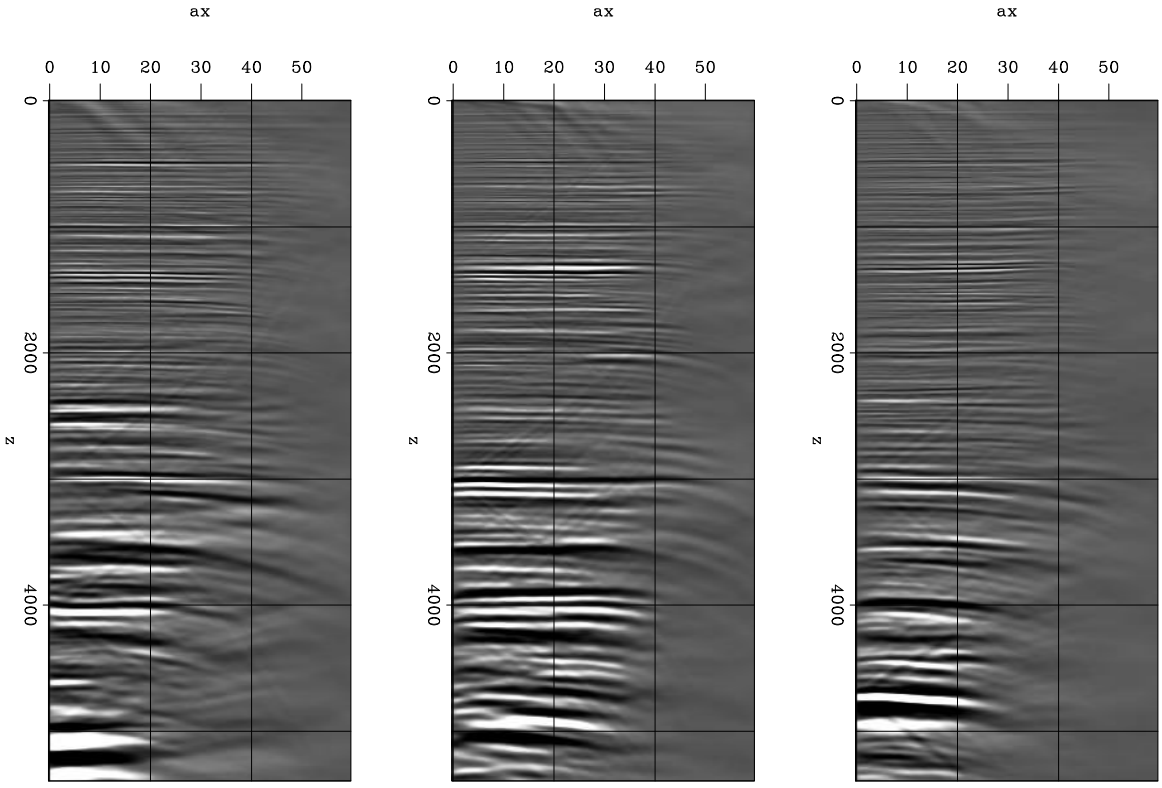


Figure 3: Angle domain common image gathers after Stolt residual migration. They corresponds to the same position as the angle gathers in Figure 2. Moreover, from left to right they also correspond to $\rho_p = 0.98$ $\rho_p = 0.99$ and $\rho_p = 1.02$ daniel1-ppsrm2 [CR,M]

be a solution.

Even though the theory of Stolt migration assumes constant velocity, Stolt residual migration can be safely applied to depth variant models, due to the dependency only on velocity ratios. This statement is also valid for multicomponent data.

REFERENCES

Clapp, R. G., 2002, Ray based tomography using residual Stolt migration: SEP–**112**, 1–14.

Rosales, D., and Biondi, B., 2002, Multicomponent data regularization: SEP–**112**, 97–108.

Rosales, D., Sava, P., and Biondi, B., 2001, Stolt residual migration for converted waves: SEP–**110**, 49–62.

Sava, P., 1999, Short note–on Stolt prestack residual migration: SEP–**100**, 151–158.

Sava, P., 2000, Variable-velocity prestack Stolt residual migration with application to a North Sea dataset: SEP–**103**, 147–157.

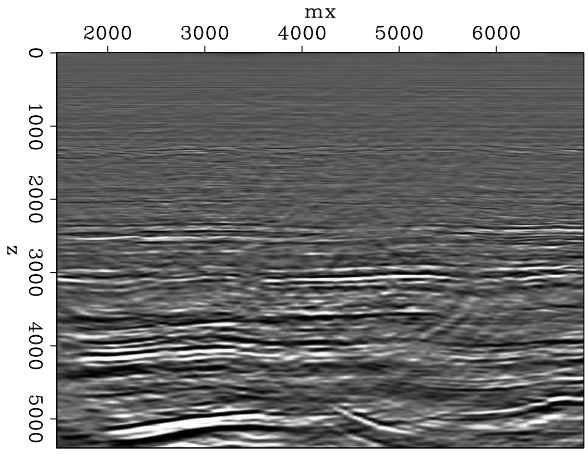


Figure 4: Migration result after updating the P velocity model with different values of ρ_p at different depth and lateral locations. daniel1-finalpp [CR,M]

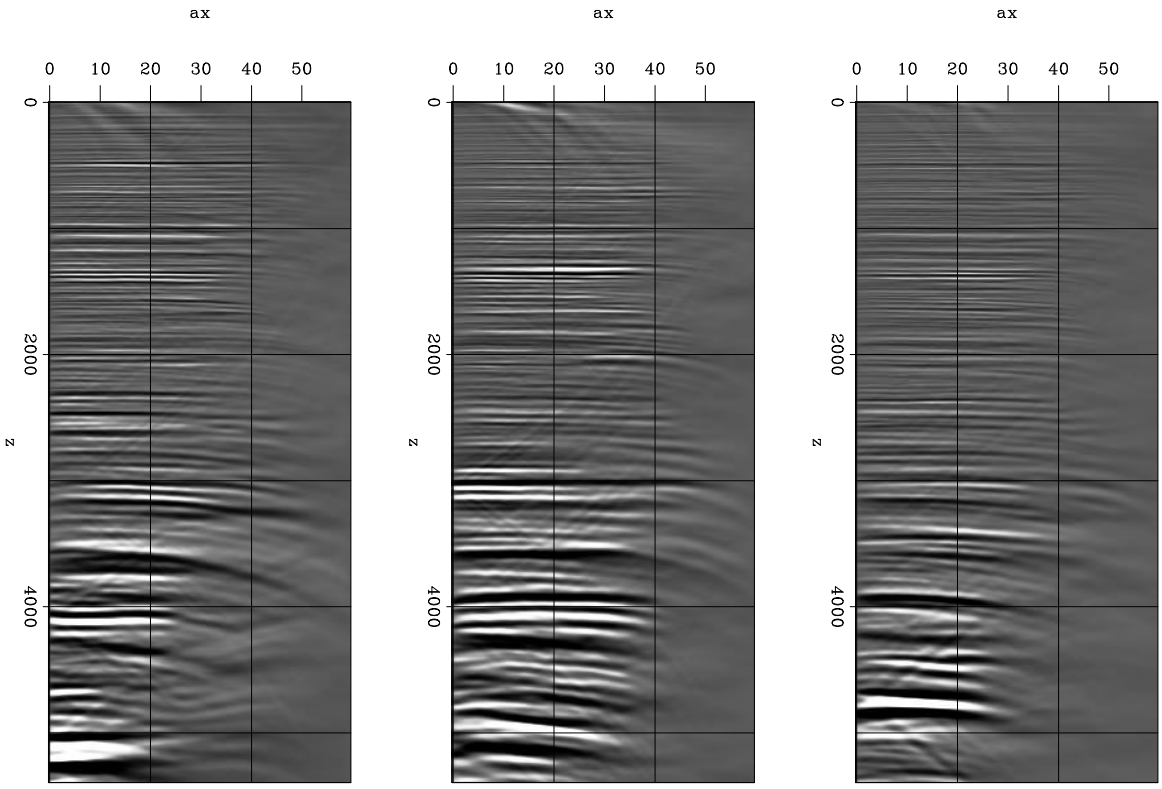


Figure 5: Angle domain common image gathers for the migration result of Figure 4 at the same position of Figure 2. daniel1-finalcag [CR,M]

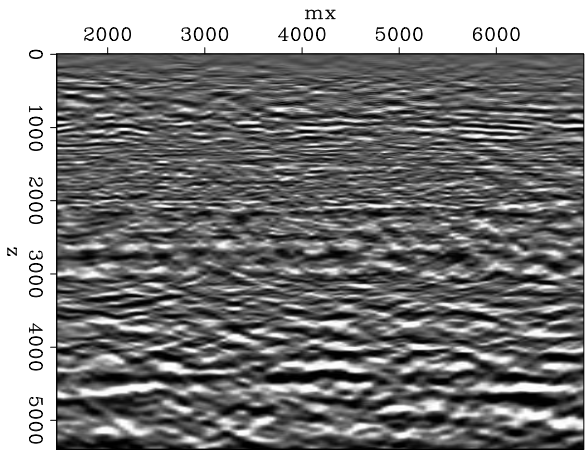


Figure 6: *PS* migration result with the updated *P* velocity model and the given *S* velocity model. daniel1-psmiga [CR]

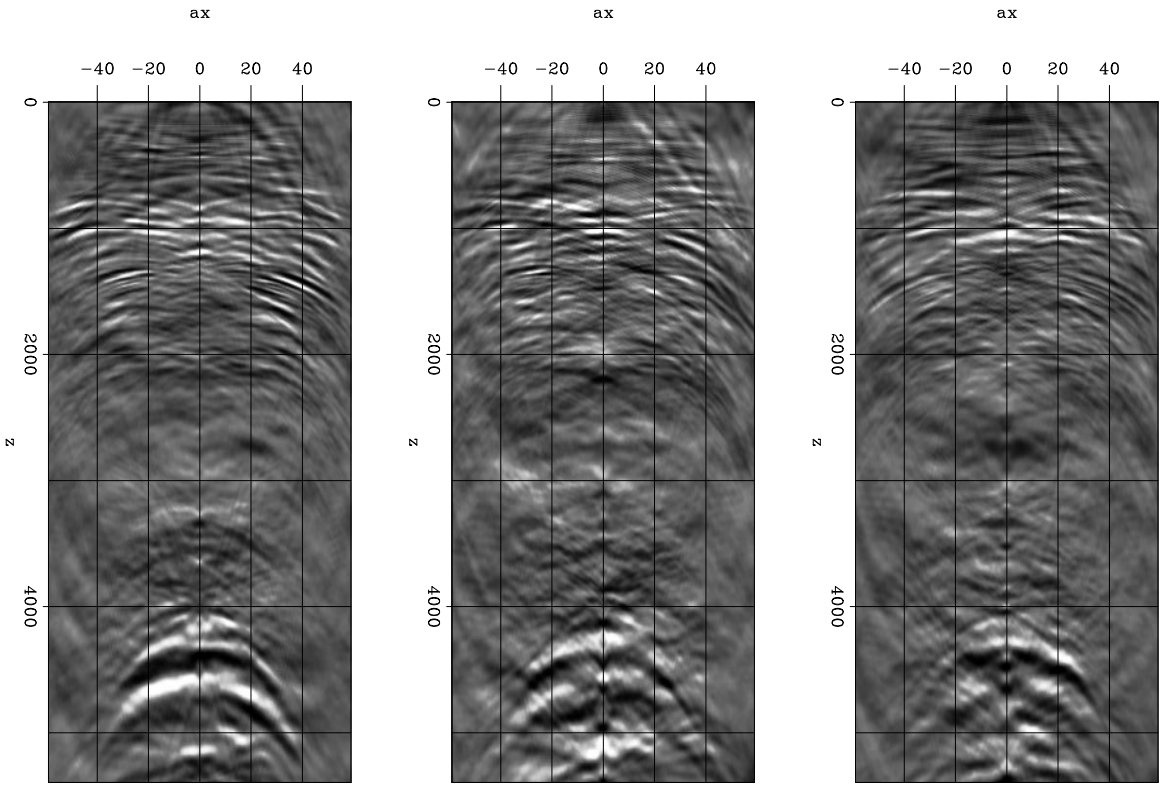


Figure 7: *PS* angle domain common image gathers. They corresponds, from left to right, to a lateral position of 2000m, 4000m and 5400m. Observe the significant residual curvature. daniel1-psmigb [CR,M]

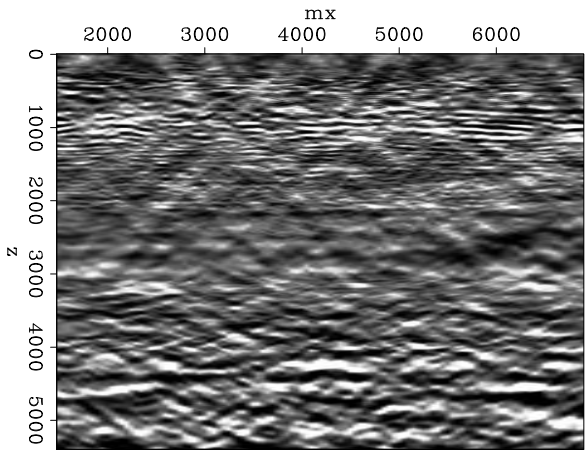


Figure 8: *PS* Stolt residual migration result daniel1-pssrm [CR,M]

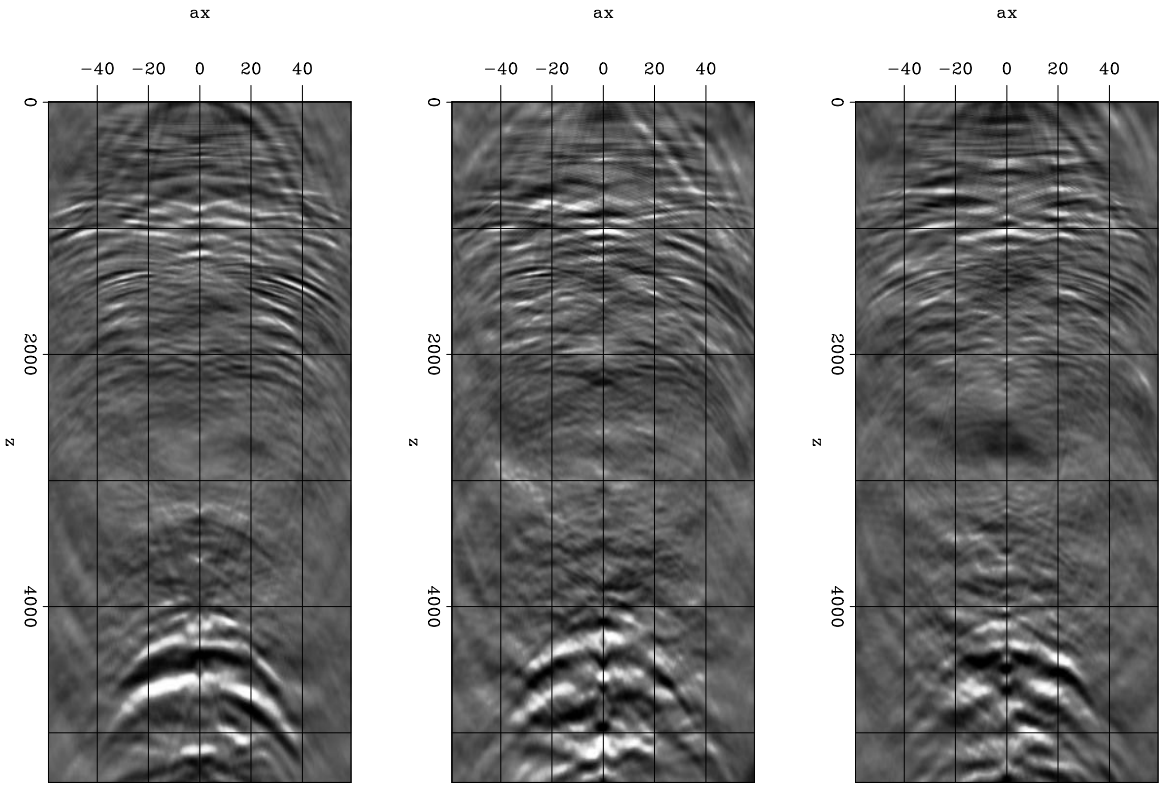


Figure 9: *PS* angle domain common image gathers after Stolt residual migration for converted waves. They are at the same location as in Figure 7 and corresponds to, from left to right, $\rho_s = 0.98$, $\rho_s = 0.99$ and $\rho_s = 1.02$. daniel1-pssrm2 [CR,M]

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 131–137

Migrating passive seismic data

Brad Artman[1]

ABSTRACT

It is possible to migrate raw passive seismic data with a modified shot-profile migration algorithm to produce a subsurface image. This skips time-intensive and space-consuming pre-processing steps as has heretofore been assumed necessary. Further, output sections are better focused and more accurately imaged using less computer time.

INTRODUCTION

Claerbout (1968) provides a one-dimensional proof that sparked the idea of imaging the subsurface without a source. By auto-correlating time series collected on the surface of the earth, he shows that one can produce the equivalent to a zero-offset time section. Zhang (1989) extends the one-dimensional proof of that conjecture, through plane-wave decomposition, to full space. Importantly, he also shows that by cross-correlating each receiver with every other, one constructs pseudo shot gathers as a function of offset. These gathers are identical to conventional shot gathers and can be treated as such throughout further processing steps.

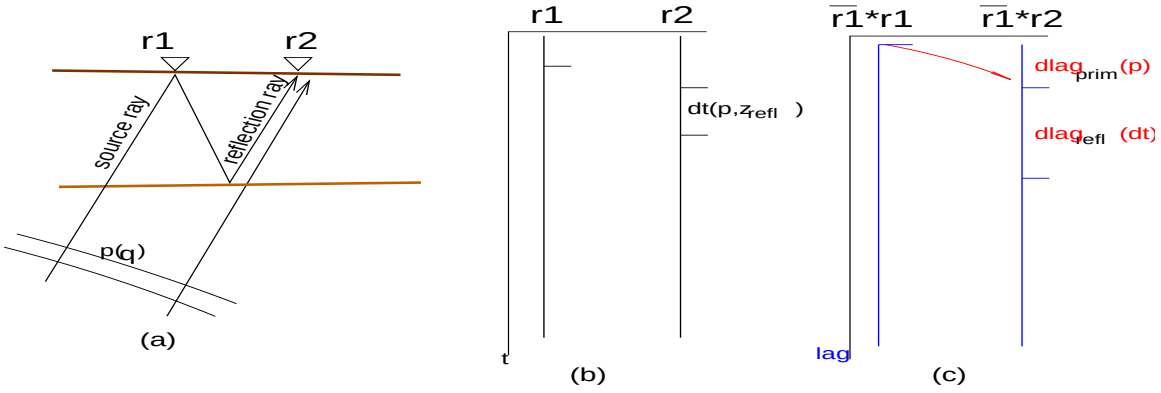


Figure 1: (a) Cartoon showing the incident plane-wave energy reflecting from the free surface and then again from a buried reflector. (b) Time-series built from arrivals depicted in left panel. (c) Correlations of signals from receivers $r1$ and $r2$. brad1-noise [NR]

The cartoon in Figure 1 shows schematically how an upcoming wave-train, $P(\theta)$, will reflect from the surface and act as a source at receiver r1 for the subsurface reflection recorded

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

at r2. An intuitive understanding of this experiment and the manufacture of pseudo shot gathers can be developed by progressing through the panels of Figure 1 from the earth model in panel (a) through the raw data recording in panel (b) to the correlated records in panel (c) where the fast-axis is now in units of correlation lag.

The correlation step has two important functions when applied to raw passive data. First, it collapses to an impulse the long source functions of the ambient sound energy recorded at each station. Second, it develops hyperbolic move-out of these impulses as a single input trace is correlated with data from stations at increasing offset. Figure 2 shows a representative gather from a synthetic passive dataset over a two-layer earth-model.

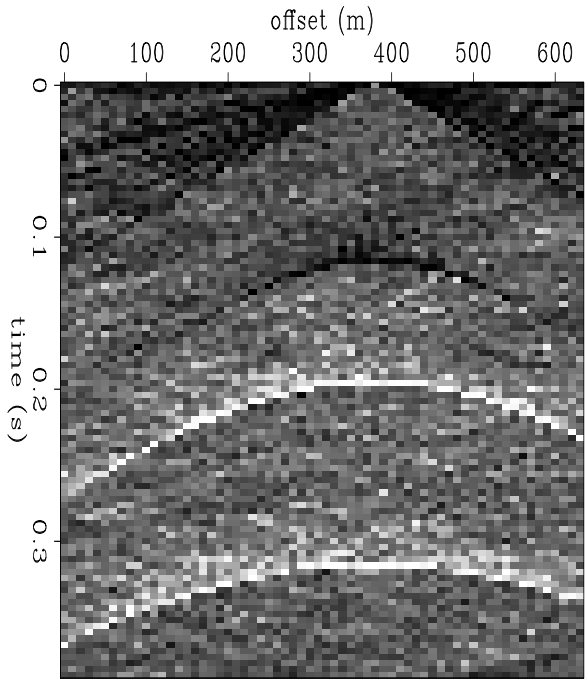


Figure 2: Representative pseudo shot gather from data cube generated by cross-correlating all traces from a modeled passive seismic dataset. The raw data trace that has been correlated with all the others serves as the source location for the gather. Notice the "virtual multiple" at 0.11 seconds. This arises due to the partial correlation of the two reflectors between themselves. The greater the velocity contrast, the less visible this event. Its zero-offset time is equal to the difference in time between the two events. brad1-shot [CR]

After the pseudo shot gathers are constructed through correlation, a conventional five-dimensional data volume is ready for any common processing flow such as sorting and migration. However, I show that it is possible to migrate the raw data directly with a shot-profile migration algorithm that has been modified to use an areal source rather than a conventional impulse or wavelet. I present comparisons of both methods performed on modeled data and comment on some of their characteristics and costs.

METHODOLOGY

The imaging condition for shot-profile migration (Claerbout, 1971) is

$$I(\mathbf{x},z) = \sum_{\omega} P^{g}(\mathbf{x},z,\omega)\overline{P^{s}(\mathbf{x},z,\omega)}, \tag{1}$$

where the image, I, is a function of surface location, $\mathbf{x}$, and depth, z, and geophone and source wavefields, $P^{g,s}$, are functions of location, depth and frequency, ω. I hypothesized that the

correlation in the imaging condition would satisfy that in the passive seismic conjecture and make calculating the correlations prior to processing unnecessary. Further, we could rely on the dispersion relation to handle the unknown phase characteristics of the ambient noise-field rather than hoping that the correlations will collapse these wave-trains into a well-behaved wavelet.

Therefore, without making the intermediate processing step of correlating all traces with each other, we can downward continue the receiver wavefield, P^g, from every location back into the earth. This means we are migrating the entire dataset as one large shot gather. Remembering the cartoon in Figure 1, we can comfortably accept the same wavefield for P^s since the source wavefield is recorded by each receiver as it reflects from the free surface. Setting $P^g = P^s$, I then migrate the data with a modified shot-profile algorithm similar to that presented in Guitton (2002).

DISCUSSION

Figure 3 shows the result of migrating the gathers manufactured through correlation in panel (a) and migrating the raw data directly in panel (b). The raw passive data was modeled over a two layer earth-model and looks like random noise.

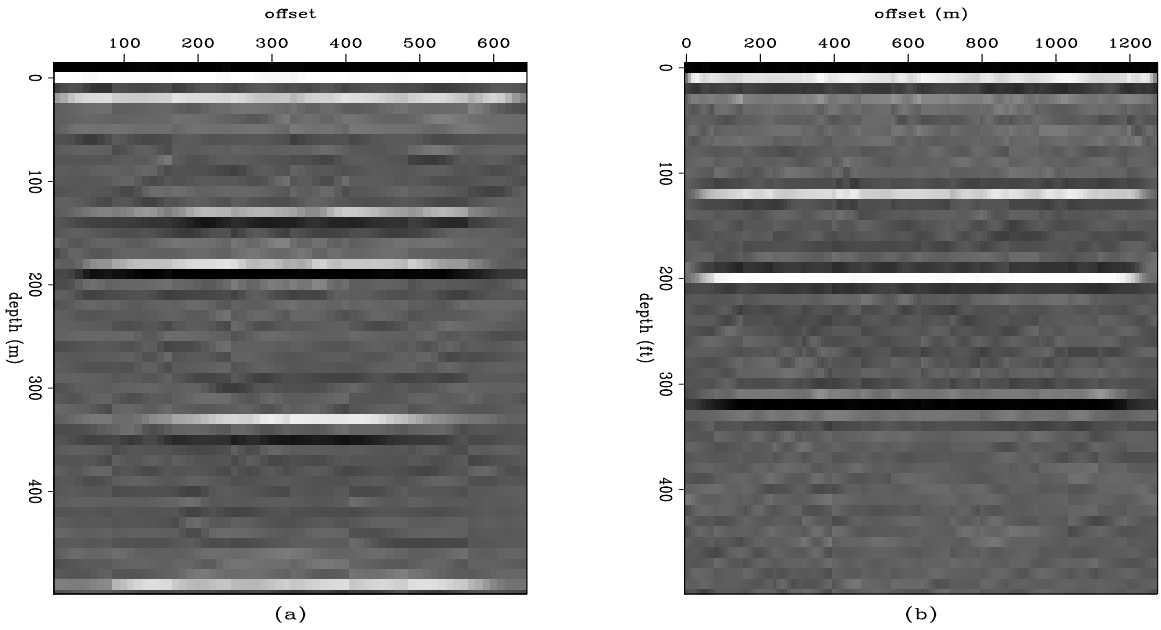


Figure 3: Left panel is result of migrating manufactured shot gathers. Right panel is the result of migrating the raw data directly. Notice the virtual multiple at 120 meters and the ramping of amplitudes of the two events from the edges to the center. The phase differences between the two panels is a result of the incorrect wavelet introduced in the migrations of the correlations. brad1-comp [CR]

In the migrated sections, we see again the presence of the virtual multiple mentioned in

the caption of Figure 2. This event is imaged at 120 meters depth which is the same as the separation between the two events in the model. This event exists because the two reflectors correlate with each other as well as the free surface. Real data may not be prone to this problem due to the saving grace of intra-bed multiples. Such an arrival would have opposite polarity from the virtual multiple and thus be destructed. Because of overly simplistic modeling code, such intra-bed multiples are not part of this test data.

Another feature of the section that deserves note is the ramping of amplitude of the reflectors from the edges to the middle of the model. This phenomenon is due to the experiment enjoying a linear, monotonic fold increase from one at the edges to half of the number of receivers at the center.

The size and cost differences for the two starting points described above for processing are significant. Making the correlation cube from the raw data squares the size of the data. However, after correlation, it is no longer necessary to maintain the extraordinarily long time series of the original data. We are free to discard all of the correlation lags computed after longest time the survey is actually interested in and only need migrate that many frequencies. By thus doing so, we shrink the data back down to about its original size. Therefore, the size of the data sets input to migration are roughly equivalent whether we consider the raw data or the correlated shot gathers. The large difference in processing time comes largely in sorting and write statements. By migrating all of the raw data as one shot gather, we enjoy operating on one entirely populated model space with only loops over depth and frequency. In contrast, the correlated data has the number of receivers equal to the number of shots to loop through, each of which populates only a small segment of the model space.

Comparisons of the time taken to migrate the two sections shown in Figure 3 show the correlated sections taking a bit more than twice as to compute with the same program. This does not include the time needed to produce the correlation volume from the raw data to use as input which makes the comparison even worse. Further, the raw data migrated section shown here was computed with a new parallel migration program that runs on our multi-node computer cluster. This architecture is well suited to the structure of real passive seismic datasets where we can expect a reasonably small model space and need to loop through a huge number of frequencies due to the multiple hours of recording.

Lastly, due to the source wavefield being completely full instead of incredibly sparse, as in conventional shot-profile migration, there is an opportunity to investigate better imaging conditions as discussed in Valenciano and Biondi (2002). This type of advanced imaging condition would also address the existence of the virtual multiples if the intra-bed multiples are of insufficient strength to cancel the multiple in real data considerations.

CONCLUSION

Migrating raw passive seismic data directly produces a cleaner, more crisp image than first correlating the input traces before migration. This new method is faster and requires less overhead, and is very efficient.

Several passive seismic datasets have been collected or acquired from various sources in the last few months. Now that a fully-parallel 3D code has been completed, I will begin processing real data examples that include shallow investigations, ocean bottom cable data, and conventional exploration geometries.

To continue to advance this effort, I am in great need of quality data acquired with as many receivers as possible. It would be very easy to collect such data by simply recording passively from receivers left in the field on a 3D acquisition overnight or during any sort of lengthy down-time. In order to best test the validity of this experiment, simple geology that provides excellent quality conventional data such as the flat limestone of west Texas would be preferred.

ACKNOWLEDGMENTS

Antoine Guitton's paper in the previous report congealed a vague hypothesis into a testable theory with his presentation of migration with an areal source. Development of the 3D parallel code used the WEI libraries of Bob Clapp and Paul Sava. Their considerable infrastructure and willingness to explain it made this product obtainable in a much shorter amount of time than could otherwise be hoped.

Partial funding of this research is provided by grant ACS PRF#37141 -AC 2 from the Petroleum Research Fund.

REFERENCES

Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response: Geophysics, **33**, no. 02, 264–269.

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

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

Valenciano, A. A., and Biondi, B., 2002, Deconvolution imaging condition for reverse-time migration: SEP–**112**, 83–96.

Zhang, L., 1989, Reflectivity estimation from passive seismic data: SEP–**60**, 85–96.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 137–141

Short Note

Deconvolving passive data

Brad Artman and Jon Claerbout[1]

INTRODUCTION

Deconvolution prior to autocorrelation processing for a passive seismic data set has the potential to ameliorate wave-parameter and azimuthal inconsistency of arriving energy during acquisition. If any particular subset of plane-wave energy dominates the passive recording sequence, full illumination of the model-space may not be achieved. Further limitation of the result could also arise from the fact that the bulk of the ambient energy recorded in the experiment will likely be ground-roll energy that does not probe the subsurface. Thus, damping over-represented energy components by convolving the data with a prediction error filter (PEF) prior to processing/migration could serve to mitigate these short-comings of the experimental design.

To address the first problem of wave-number and azimuthal inconsistency, Artman (2002) suggests a trace balancing scheme after transformation of the raw data into (ω, p)-space. Convolution with a PEF instead will also produce this result.

EXPERIMENT

To explore the feasibility of this idea, we used the solar data set that has been seen in several previous reports. The frequency content of this data shows two distinct modes as seen in Figure 1. The previous results generated by autocorrelation processing (Rickett and Claerbout, 1999) used a low-cut version of the data that removes the energy below 2 mHz. Figure 2 is a comparison of three different inputs to the autocorrelation processing algorithm. The right panel uses the raw data. The center panel uses a low-cut version of the data. The left panel is the output after 1-dimensional deconvolution. Unfortunately, a small amount of DC noise has survived the deconvolution process. Because the solar data is not quite as stationary as had been hoped, estimating the PEF on too small an area resulted in an output spectrum that is not white in all locations. Some color left in the frequency content of the result would seem acceptable; however, the low-frequency contribution to the result presents a problem in the auto-correlation processing.

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

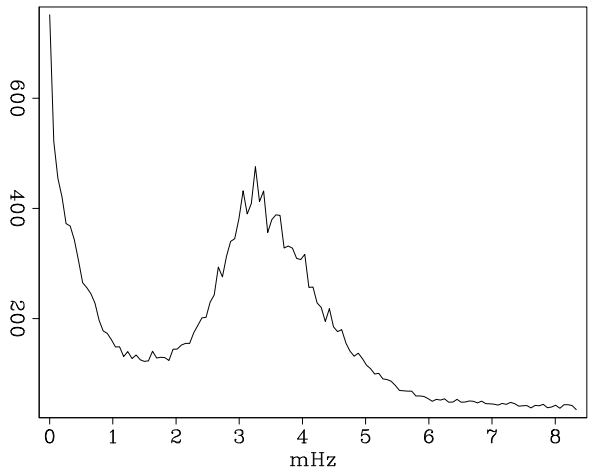


Figure 1: Frequency content of the passive solar seismic data set. Data are sampled in Kilo-seconds yielding milli-Hertz frequencies. brad2-spec [CR]

To better combat this near-D.C. component of the data, all of the traces were passed through a 1D gradient operator. Having thus removed the low-frequency component of the data, the PEF estimation and convolution process was performed again followed by auto-correlation processing. Figure 3 shows the result of this processing scheme. The result is better focused, crisper and more pleasing than the results shown in Rickett and Claerbout (1999) generated by low-cut filtering and auto-correlation processing.

The first in-line section of the result of this processing chain is shown in Figure 3 and seems to show three faint events that have not been previously identified. Unfortunately, further evidence to corroborate them as real events have not been fruitful. Perpendicular sections do not reveal similar events. Time slices do not show circular horizon intersections. Finally, an azimuthal stack around the central trace was calculated. Figure 4 shows this result with no indication of the earlier events.

CONCLUSION

Deconvolution of a passive seismic dataset has been shown to produce sharper, more crisp output. By balancing the energy recorded in an experiment arriving around all azimuths and from all incidence angles is an important first step in the passive seismic imaging experiment.

One-dimensional deconvolution of the solar passive seismic dataset proves to be a quick and advantageous step prior to auto-correlation processing to produce a sharper result. The assumption that the solar data are approximately stationary, however, is flawed and results in inadequate representation of the entire body of the data when a PEF is estimated on a small part of the data. By operating instead on the gradient of the raw data, this problem is greatly reduced as the major problematic remnant in the raw data deconvolution was a large low-frequency component.

Processing the solar data in this manner may have uncovered a previously undiscovered and possibly directionally propagating event. Conversation with the solar physicist are underway in hopes of identifying what type of physical phenomenon these could be. Two-

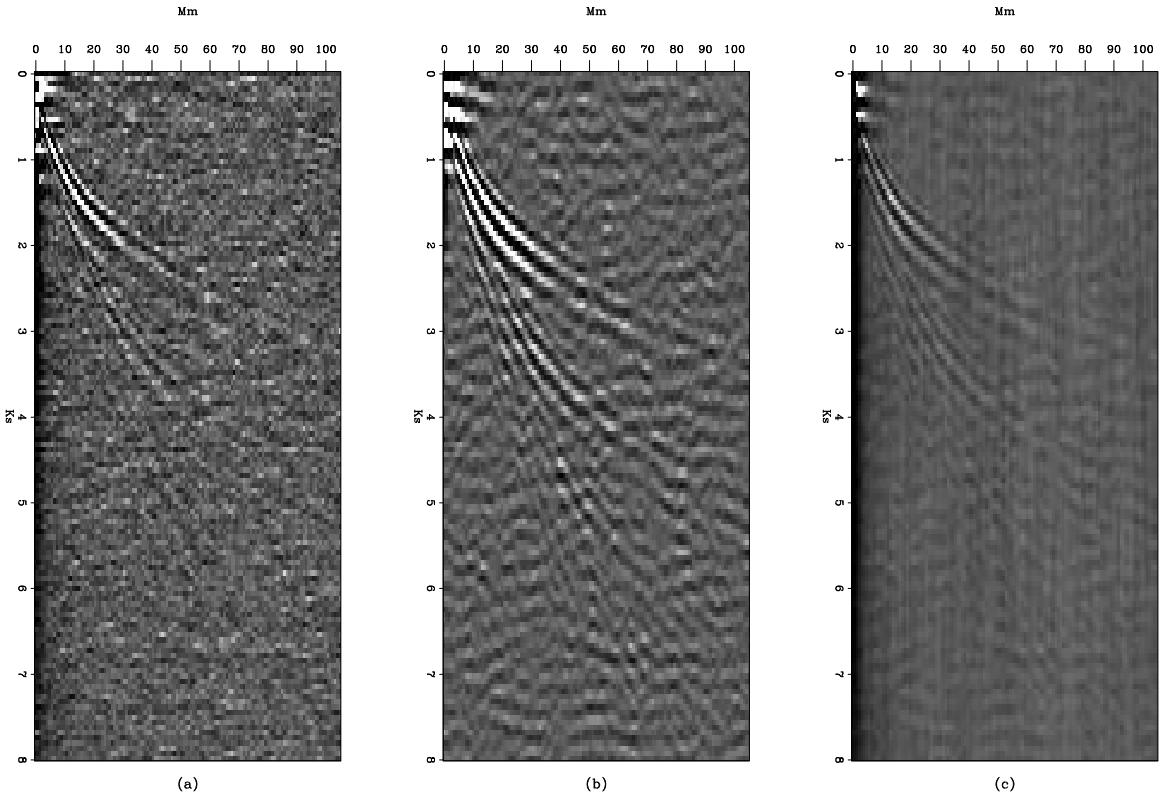


Figure 2: Left panel is autocorrelation after deconvolution. Center panel is processing after low-cut filter. Right panel is produced using the raw data. brad2-three [CR]

dimensional deconvolution and work on a larger, longer similar data set may prove fruitful in the near future.

ACKNOWLEDGMENTS

We would like to thank the SOI/MDI project at the Stanford Physics Department for donating this dataset and one to come. Thanks also to James Rickett on whose shoulders this work progresses.

REFERENCES

Artman, B., 2002, Coherent noise in the passive imaging experiment: SEP–**111**, 379–383.

Rickett, J., and Claerbout, J., 1999, Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring: SEP–**100**, 171–180.

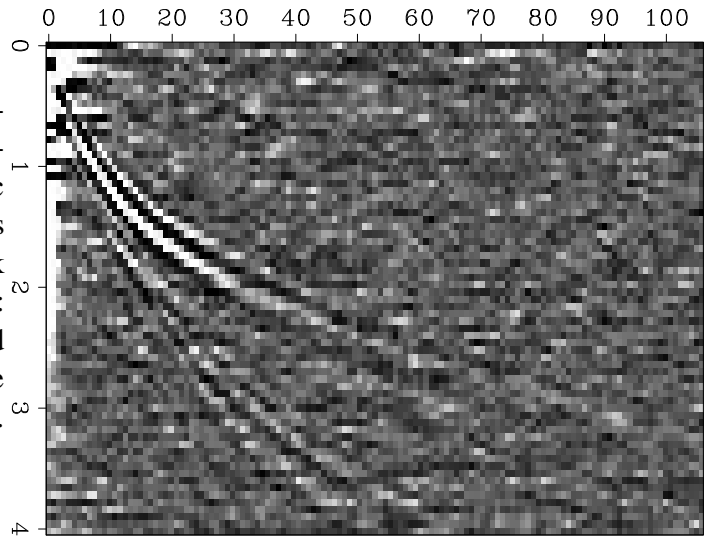


Figure 3: After 1D gradient, deconvolution, and auto-correlation processing the solar passive seismic data may reveal faint new events on this in-line section. Look carefully for 3 dark linear events: from $(0Mm, 0ks)$, $(10Mm, 0ks)$, and $(40Mm, 0Ks)$. The velocity of these events is approximately $50,000m/s$. brad2-newevent [CR]

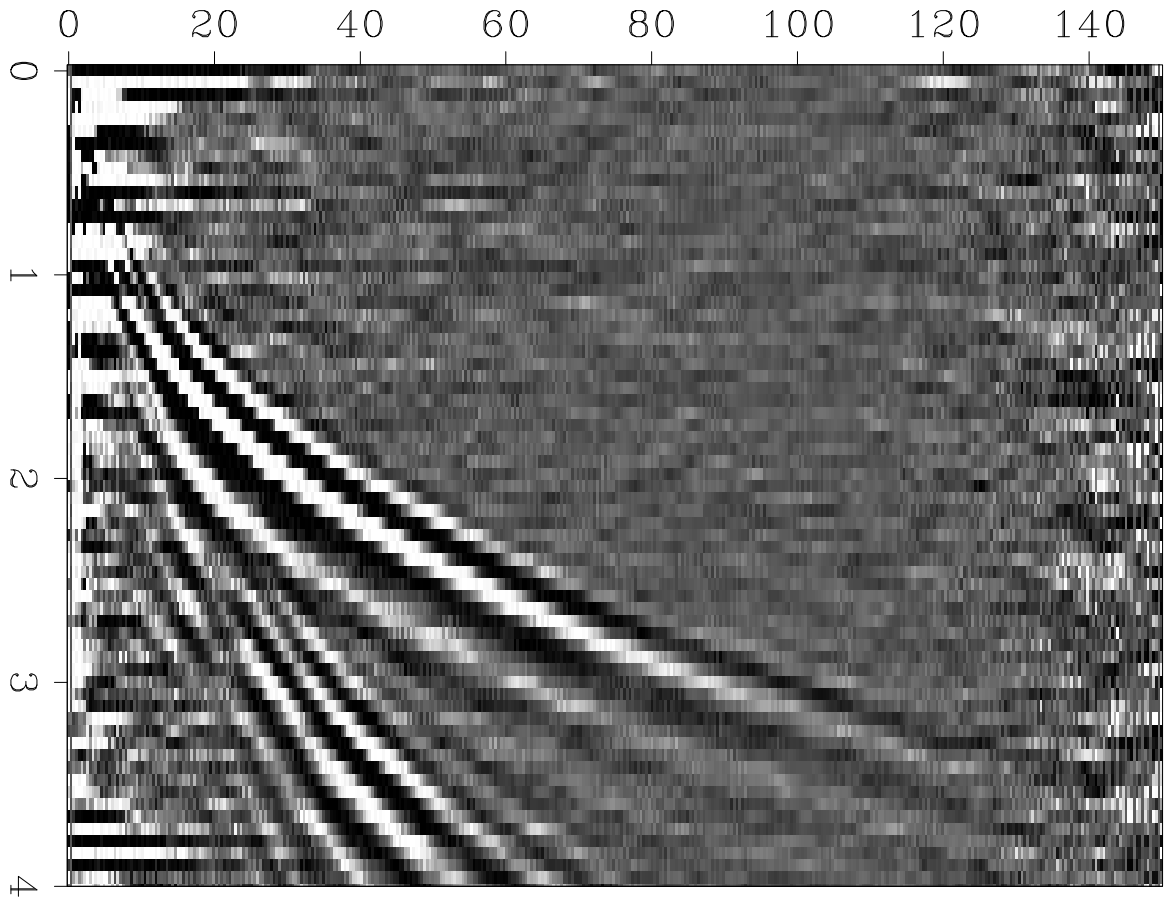


Figure 4: Radial stack of previous data to see if the "events" might stack into reality. No further evidence of the events. brad2-radstack [CR]

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 141–151

Flattening without picking

Jesse Lomask and Jon Claerbout[1]

ABSTRACT

We introduce an analytical method for integrating dip information to flatten uninterpreted seismic data. First, dips are calculated over the entire seismic volume. The dip is then integrated in the Fourier domain, returning for each sample a time shift to a flat datum. Then each sample is shifted in the seismic data to remove all structural folding deformation in a single non-interpretive step. Using the Fourier domain makes it a quick process but requires that the boundaries are periodic. This method does not yet properly handle faults because of their discontinuous nature, but is presently very effective at removing warping and folding.

INTRODUCTION

A commonly used interpretation technique is to flatten data on horizons. This removes structure and allows the interpreter to see geological features as they were laid down. For instance, after flattening the seismic data, an interpreter can see an entire flood plain complete with meandering channels in one image.

Previously, in order to flatten seismic data, a horizon needed to be interpreted. If the structure was changing then many horizons needed to be interpreted. Here we propose a method for automatically flattening entire 3D seismic cubes without any interpretation at all. Our method involves first calculating dips everywhere in the data using a dip estimation technique described in Claerbout (1992). The local dips are resolved into a local travel time via a least squares problem that we solve in the Fourier domain. Then the data is shifted according to the travel times to output a flattened volume.

In this paper, we review the method for calculating dips and describe, in detail, the dip integration method. Then we show the results of several test cases. The first test case is a 3D synthetic data set with planes dipping at a single dip everywhere. The second test case is also a 3D synthetic data set, but has curved horizons. Finally, we apply this method to flatten horizons warped by a salt piercement in a real 3D seismic data set from the Gulf of Mexico.

This method successfully flattens the synthetic test data sets and removes a lot of deformation from the real test data set. These results are encouraging and invite more testing with more complicated models. The ability of this method to flatten data with faults still needs further development.

[1]**email:** lomask@sep.Stanford.EDU,jon@sep.Stanford.EDU

METHODOLOGY

First we calculate the dip. Dip can be easily calculated using a plane-wave destructor as described in Claerbout (1992).

For the dip in the x direction of a seismic cube with a wave field represented by $u(x,y,t)$, at each sample we calculate:

$$p_x = -\frac{x' * t'}{t' * t'} \tag{1}$$

where x' is the $\partial u/\partial x$ taken on a mesh in (x,t) and t' is $\partial u/\partial t$. Because we are calculating a different dip at each sample, it is necessary to smooth the dips. We apply a triangle filter to both the numerator and denominator of equation (1). Presently, in calculating p_x, we smooth along the x-axis and t-axis. However, a more robust approach would be to smooth along the x-axis, t-axis, and y-axis.

Our main objective is to find an absolute time (t) at each sample in the seismic data cube. Because the dip can be thought of as the gradient (∇), the dip in the x direction (p_x) is the x component of the gradient. Similarly, the dip in the y direction (p_y) is the y component of the gradient. Using our integration method described below, we first apply the divergence (∇') to the gradient. Then we convert to Fourier space where we integrate twice by dividing by the Laplacian. Then we convert back to the time domain. The resulting t can be thought of as the absolute time for each point in the data.

Beginning with our input dip data:

$$\nabla t \quad = \quad \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial t}\right) t \tag{2}$$

where $\frac{\partial}{\partial x} = p_x$, $\frac{\partial}{\partial y} = p_y$, and $\frac{\partial}{\partial t}$ is all ones for smoothness in time (explained below).

The analytical solution is found with:

$$t \quad \approx \quad FFT^{-1}\left[\frac{FFT\left[\nabla' \nabla t\right]}{-Z_x^{-1} - Z_y^{-1} - Z_t^{-1} + 6 - Z_x - Z_y - Z_t}\right] \tag{3}$$

where $Z_x = e^{iw\Delta x}$, $Z_y = e^{iw\Delta y}$, $Z_t = e^{iw\Delta t}$ and FFT is the 3D Fourier transform.

The denominator is the Z-transform of the 3D Laplacian. The zero frequency term of the Z-transform of the denominator is neglected. This means that the resulting surface in space will have an unknown constant shift applied to it. However, by adding the t dimension and assuming the gradient in the t direction to be all ones, we are insuring that the integrated time varies smoothly in the t direction.

Integrating in three dimensions enforces vertical smoothness. The dip in the t direction is all ones. This can be thought of intuitively as imagining that the dip in the x direction is the derivative of x with respect to t. So dip in the t direction is the derivative of t with respect to t, therefore it is always one. By integrating in 3D, we prevent our method from swapping sample positions in time.

Boundaries

Artifacts were created by this method when the boundaries of the result of the divergence were not periodic. One way to solve this problem is to make mirror images of the input dip data. Mirroring, which is basically replicating and time reversing data, requires increasing the data size by a factor of four in 2D and a factor of eight in 3D. Another way that requires less memory, is to define a new gradient operator that is periodic.

A periodic gradient operator can most easily be explained with a one dimensional example. A non-periodic gradient operator differences all of the samples in a one dimensional array. The periodic gradient operator does the same but also it differences the first and last samples. In equation (3), the application of the divergence (∇') with a periodic gradient operator outputs a periodic result that when Fourier transformed will better match the periodic denominator.

If we were to use the periodic gradient operator without mirrors in our formulation then in calculating the dip, we would need to know the dip from one side of the image to the other. This eliminates the periodic gradient as a solution to our problem in the x and y directions but not the t direction. In the t direction, where the gradient is one everywhere, we know the gradient from one end to the other. It is equal to the total time. As a result, we are using the periodic gradient operator in all three dimensions but only need mirrors in the x and y dimensions. This increases our data size by a factor of only four, rather than eight, even though it is in 3D.

TEST CASES

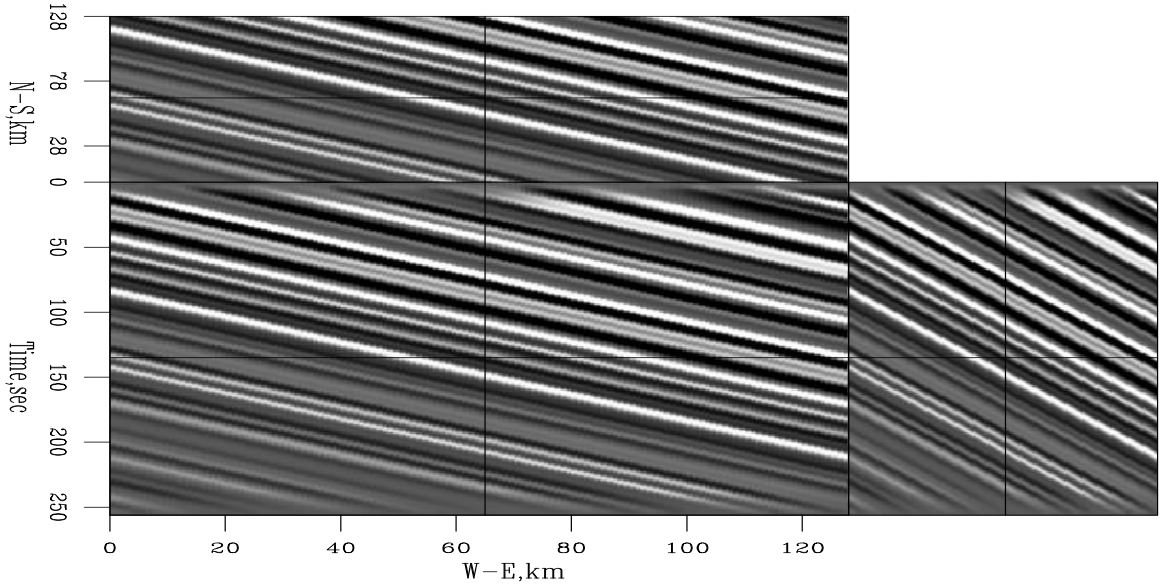


Figure 1: Test case 1. Dipping planes. jesse1-plane3D [ER]

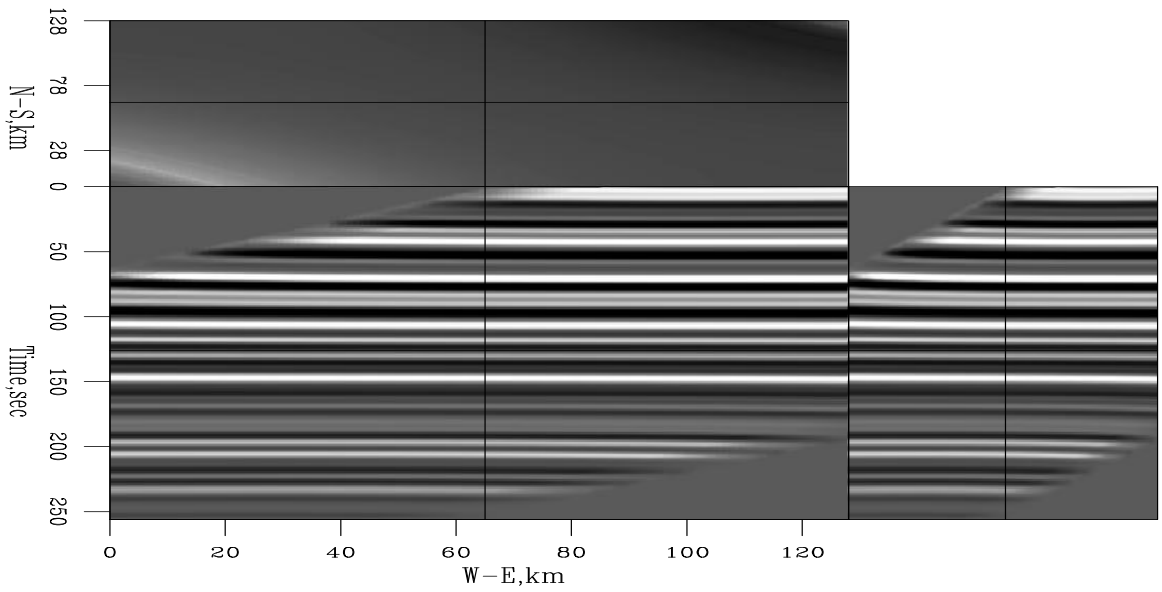


Figure 2: Result of flattening data in Figure 1. jesse1-plane3D.3Dflat [ER]

Dipping planes

The 3D synthetic data set in Figure 1 consists of a single synthetic seismic trace that is replicated and delayed so that the dip in both the x and y directions is unity. Running a dip estimator as in equation (1) will ideally result in dips of one in each direction.

The results of the flattening method of the data in Figure 1 are shown in Figure 2. Notice that the dipping planes are now flat and that the time slice on top is all one gray tone, indicating that it is flat.

Curved horizons

Figure 3 shows our next test case. The surfaces are curved upwards along the x-axis. The results of the flattening are shown in Figure 4. The method has successfully flattened the data. The slight undulations that look like interpolation errors are present in the original model in Figure 3.

Real 3D data

Figure 5 is a real 3D data cube from the Gulf of Mexico provided by Chevron. It consists of almost flat horizons that have been warped up around a salt piercement. Numerous channels can be seen in time slices. In the time slice at the top of Figure 5 a channel can be seen snaking across along the south side.

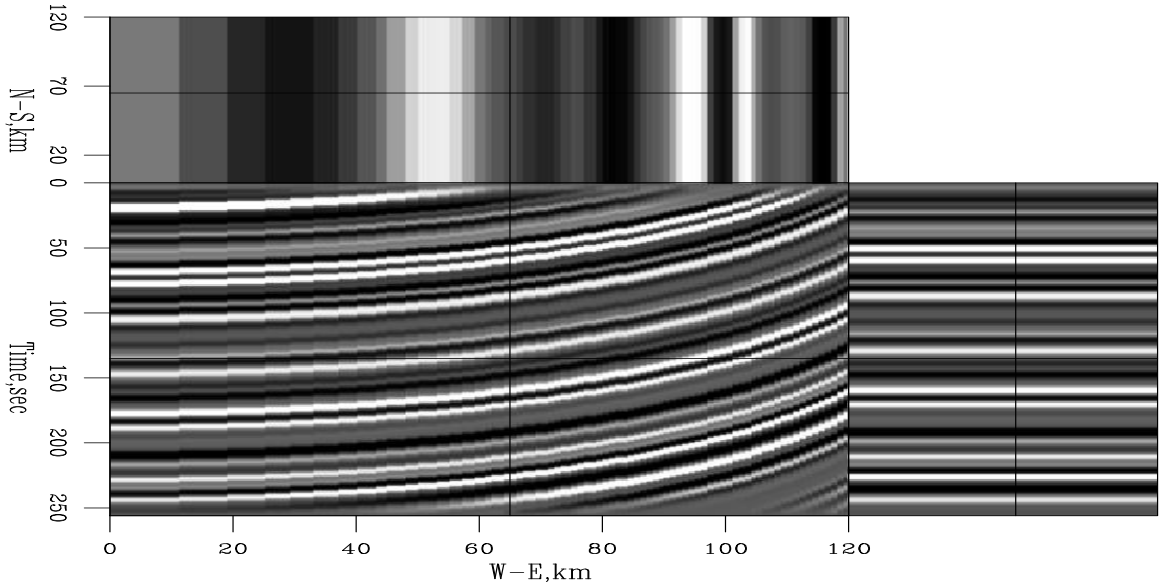


Figure 3: Test case 2. Curved horizons. jesse1-ski_jmp [ER]

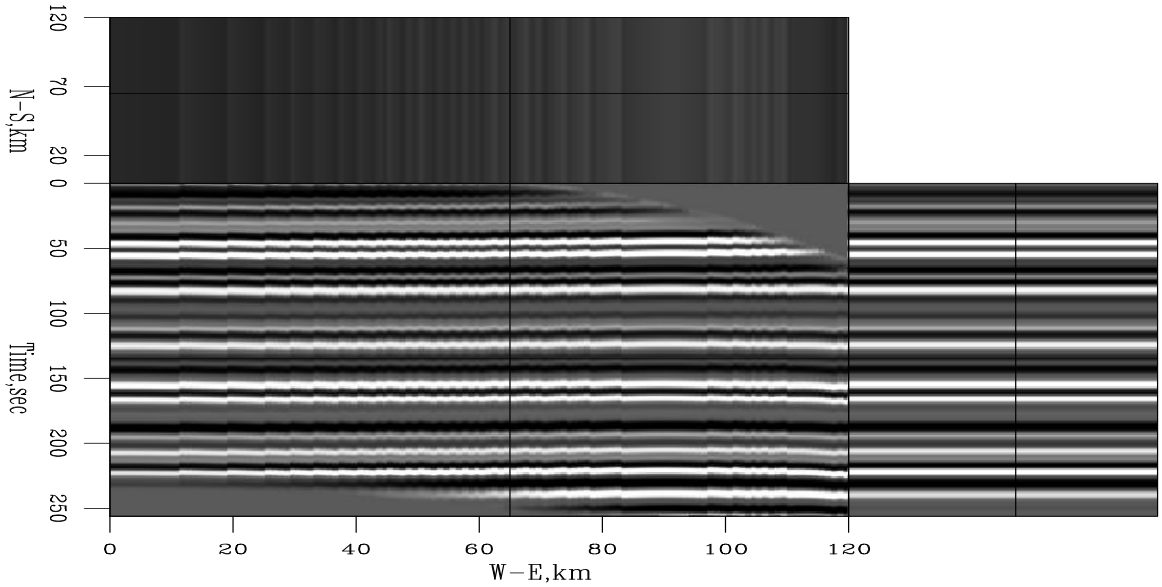


Figure 4: Result of flattening data in Figure 3. jesse1-ski_jmp.3Dflat [ER]

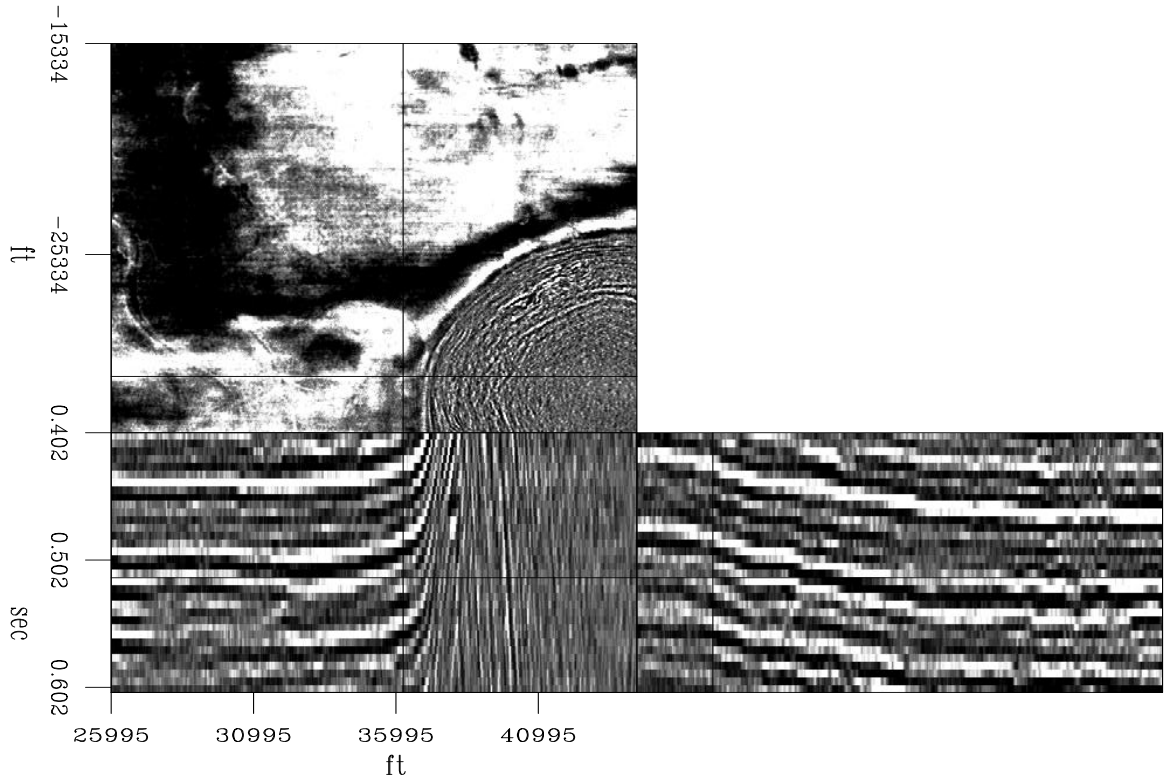


Figure 5: Test case 3. Chevron Gulf of Mexico data. jesse1-chev [ER]

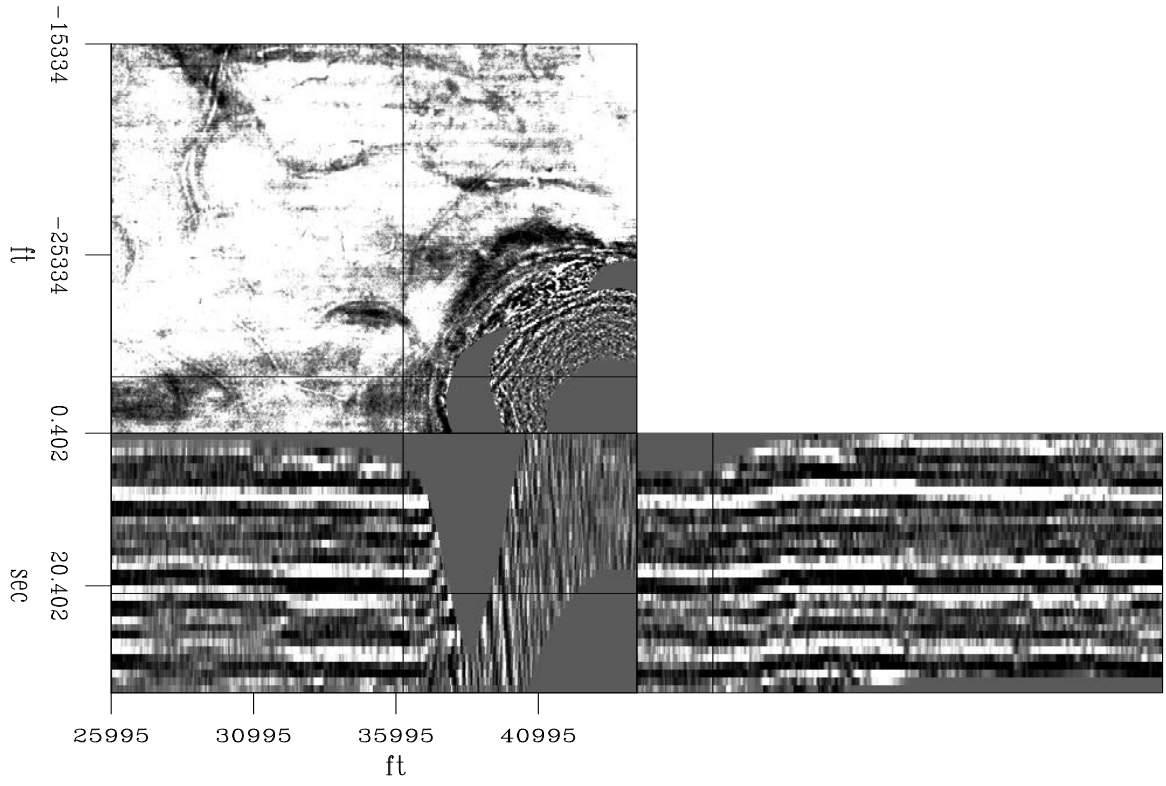


Figure 6: Result of flattening data in Figure 5. jesse1-chev.3Dflat [ER]

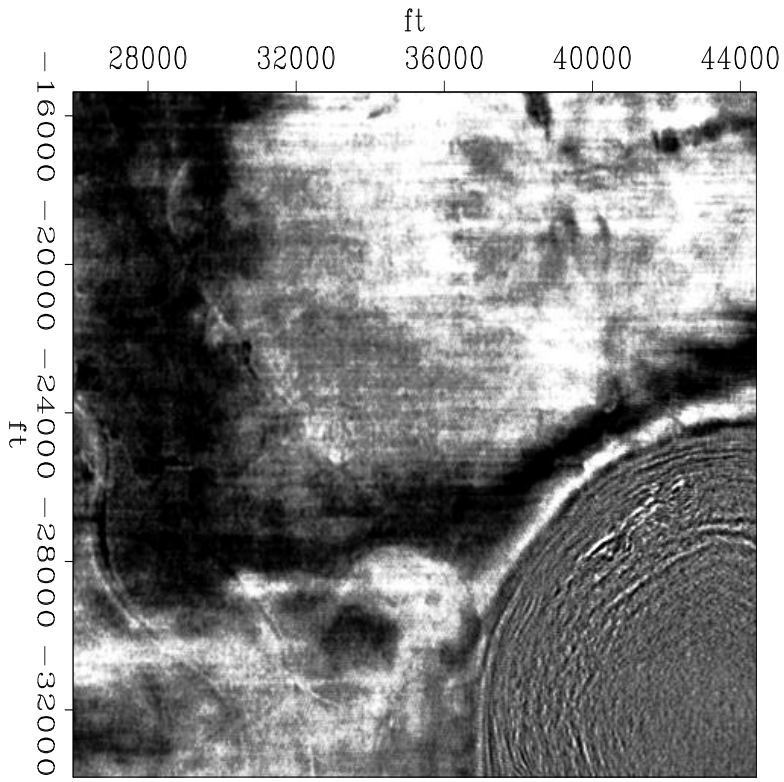


Figure 7: Time slice from Unflattened Chevron Gulf of Mexico data. jesse1-chev_ts [ER]

Figure 6 shows the flatted output of the Chevron data. Notice that the horizons are flatter than those of Figure 5.

Figure 7 is an unflattened time slice from the Chevron data and Figure 8 is a flattened time slice (horizon slice) from approximately the same place in the data. Notice Figure 8 does not have the low frequency banding that Figure 7 has. Also notice that the salt dome appears to be smaller in Figure 8. This indicates that the layers warped up by the salt have been made flatter.

Figure 9 compares an east-west section from the Chevron data to its flattened counterpart. The left side of the flattened section is clearly flatter that the input above. However, notice that the flattened image doesn't do a great job where the beds are dipping up steeply. This could be a result of poor dip estimation and definitely warrants further investigation.

CONCLUSIONS

Our method of resolving local dips into time shifts has effectively flattened seismic in our test cases.

Our use of the 3D Fourier transform may not be necessary. We maybe able to integrate the dips in 2D. This would make this method capable of handling large data sets easily. As

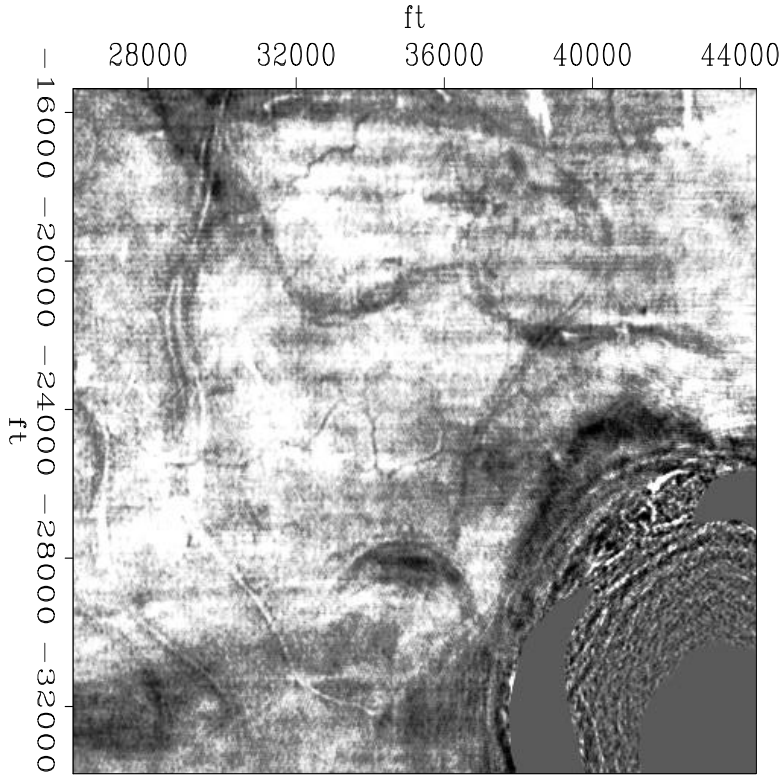


Figure 8: Time slice from Flattened Chevron Gulf of Mexico data. jesse1-chev.3Dflat_ts [ER]

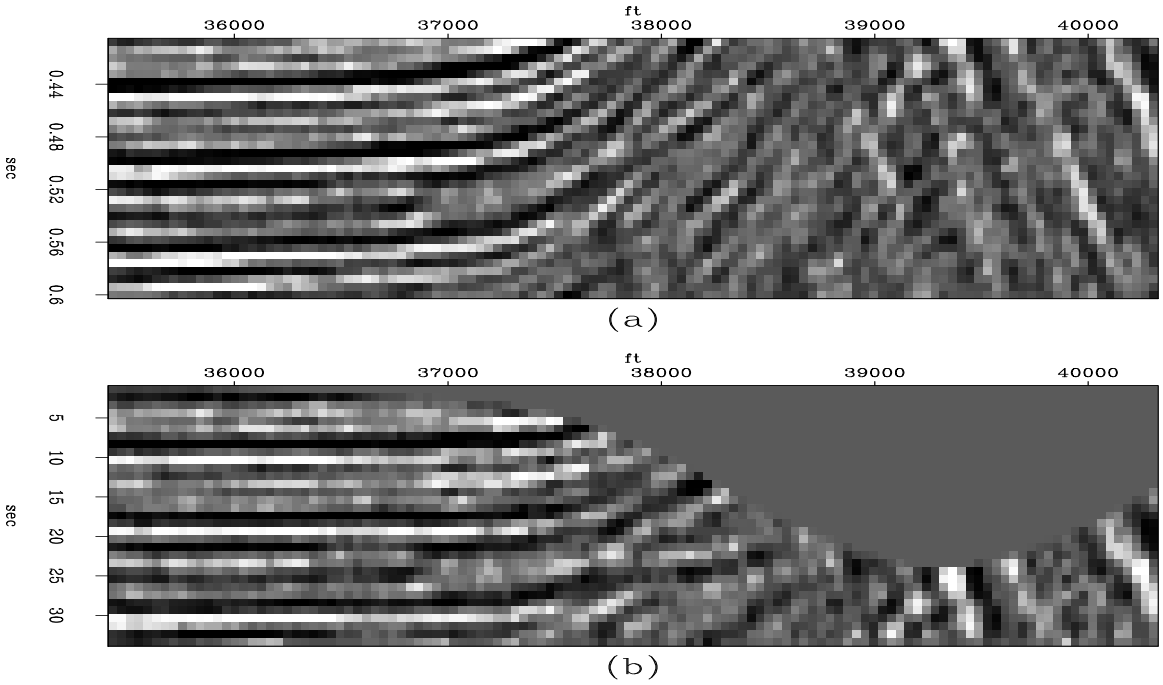


Figure 9: East west sections from Chevron data. (a) Unflattened. (b) Flattened. jesse1-chev.section [ER]

mentioned earlier, we can smooth both the numerator and denominator of the dip calculation in equation (1) along all three axis. This could possibly eliminate the need for integrating in the t direction by properly smoothing the dip calculation.

The ability of this method to work with data that has pinch outs and faults still needs to be looked at. A local dip estimator will estimate incorrect dips at faults. Compounding the problem, our dip integration method will try to honor those incorrect dips. Once in the Fourier domain, it will be very hard to correct this problem.

Overall, the results of this method are very encouraging. The ability to flatten data could be a powerful tool in automating interpretation in general. There could be many processing applications as well, such as flattening gathers.

ACKNOWLEDGMENTS

We would like to thank ChevronTexaco for the Gulf of Mexico data.

REFERENCES

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

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 151–161

Combined inversion: preconditioning with regularization

Marie L. Clapp[1]

ABSTRACT

Iterative inversion schemes are becoming more common in seismic processing. The high cost of the operators generally used in these inversion schemes makes it very important to minimize the number of iterations needed to obtain a good model. In complex environments, inversion schemes can be improved by styling the model through regularization or preconditioning. At early iterations, regularization provides a result that has a frequency content comparable to that of the "ideal" model. Preconditioning defines a solution at every model point at earlier iterations than regularization. An "improved" model should combine these two characteristics. This paper examines a scheme that uses the result of preconditioned inversion as an initial model for regularized inversion. I show that this scheme allows us to obtain an improved model in fewer iterations than would be needed for preconditioned inversion or regularized inversion alone.

INTRODUCTION

As the search for oil concentrates on ever more complicated areas of the subsurface, we find ourselves needing to balance the benefits of obtaining a better image with the cost of obtaining that better image. To obtain an ideal image, we would have to use an imaging operator that is the inverse of the physical operator propagating our seismic signal into the ground. However, imaging operators such as migration are adjoints rather than inverses (Claerbout, 1995), so in complex areas the resulting image may not be as good as it could be. Unfortunately, finding an operator that is an inverse in complex areas is almost impossible, so we generally approximate the inverse through a process like least-squares inversion (Nemeth et al., 1999; Duquet and Marfurt, 1999; Ronen and Liner, 2000). When using such an iterative technique, the result of iterating to convergence can be thought of as the "ideal" model.

Iterative inversion schemes often have trouble with problems that are unstable or where the mapping operator has a null space (Claerbout, 1991). These issues can be overcome by regularizing the problem (Tikhonov and Arsenin, 1977; Harlan, 1986; Fomel, 1997). However, our regularization operators, which are usually roughening operators, tend to be small. Their influence at any single iteration is limited in range. When our mapping operator has large areas that do not correspond to any data locations this can be especially troublesome. A solution to this problem is to perform a change of variables, turning it into a preconditioned problem

[1]**email:** marie @ sep.Stanford.edu
The author has recently changed her name from Marie Prucha to Marie Clapp.

(Fomel et al., 1997). Using the helix transform (Claerbout, 1998), we can apply the inverse of our small regularization operator, which will be a smoothing operator, whose influence extends a large distance. The advantage of this approach is that we quickly define our solution at all model points. The disadvantage of this approach is that our preconditioning operator dominates early iterations, creating a model that is often too low in frequency. What we ideally would like is a process where the solution is defined everywhere without the reduction in frequency content.

In Prucha and Biondi (2002), we presented a scheme that met these requirements by using the result of the preconditioned inversion as an initial model for a regularized inversion. From that example, and for the purposes of this paper, I will define an "improved" model as one that has a solution defined at every point and has a frequency content comparable to that of the "ideal" model. The combined inversion using preconditioning and regularization allows me to obtain an improved model with fewer iterations than would be needed using preconditioning or regularization alone. In this paper, I will take a closer look at the process of combined inversion.

In order to efficiently examine combined inversion with preconditioning and regularization (CIPR), this paper solves an interpolation problem which is much simpler than the imaging problem in Prucha and Biondi (2002). I will begin by explaining the constructed problem and the operator that is used for interpolation. Then I will present and discuss the results. Finally, I will explain my future plans for this combined inversion scheme.

CONSTRUCTING AN INTERPOLATION PROBLEM

The operators

In order to examine the results of preconditioned inversion, regularized inversion, and my proposed CIPR, I needed a problem that was easier to understand than that shown in Prucha and Biondi (2002). I am concerned with two issues: frequency content and solutions at every model point. To address the first issue, I chose to make my inversion operator a "smoother" that causes the model to have a higher frequency content than the data. This can be expressed as:

$$\mathbf{d} \approx \mathbf{Sm} \tag{1}$$

where $\mathbf{d}$ is the data, $\mathbf{m}$ is the model, and $\mathbf{S}$ is a smoothing operator that maps the average of 5 vertical points in the model to one point in the data. Since the model should be high frequency, the effects of the preconditioned inversion should be quite obvious.

Given such a simple inversion operator, creating a need for regularization or preconditioning requires that I cause the model created by inversion (fitting goal (1)) to have points that do not have solutions defined by the inversion operator. I chose to do this by introducing a masking operator $\mathbf{W}$. The combined operator $\mathbf{WS}$ will now have a null space where $\mathbf{W} = \mathbf{0}$. This changes my fitting goal to:

$$\mathbf{d} \approx \mathbf{WSm}. \tag{2}$$

To interpolate the model in the areas affected by the null space, I add a second fitting goal to fitting goal (2):

$$\begin{aligned} \mathbf{d} &\approx \mathbf{WSm} \\ \mathbf{0} &\approx \epsilon\mathbf{Am} \end{aligned} \tag{3}$$

where the new operator, **A**, is a regularization operator. I have chosen to make **A** a steering filter (Clapp et al., 1997; Clapp, 2001) generated as described in Prucha et al. (2000, 2001). Briefly, a steering filter consists of dip penalty filters at every model point, meaning that it is a non-stationary roughening operator that acts over short distances. To precondition this problem, I perform a change of variables to replace the model **m** with the preconditioned variable **p**:

$$\mathbf{m} = \mathbf{A}^{-1}\mathbf{p}. \tag{4}$$

Applying this to fitting goals (3) results in a new set of fitting goals:

$$\begin{aligned} \mathbf{d} &\approx \mathbf{WSA}^{-1}\mathbf{p} \\ \mathbf{0} &\approx \epsilon\mathbf{p}. \end{aligned} \tag{5}$$

The inverse of the steering filter ($\mathbf{A}^{-1}$) is applied using the helix transform. The inverse operator will be a smoothing operator that will act over a much larger distance than **A**.

The data

Given the operators I have chosen to use in this experiment, selecting data to test is straightforward. I need data that will result in a model that requires interpolation and will make differences in frequency content of various results obvious. Since the regularization operator is a steering operator, the data can have varying dips. To meet these simple requirements, I chose to take a 2-D slice from the familiar "qdome" model (Claerbout, 1995). The masking operator **W** contains enough zeros to defeat the inversion operator, making the regularization operator necessary. Figure 1 shows the data multiplied by the masking operator (**Wd**) I used for this experiment. I am displaying it this way to make comparison with the inversion results simpler. Figure 1 also shows the "ideal" model that would be obtained if **W** was simply an identity operator.

RESULTS

The first experiments I ran were to simply test the result of the regularized inversion (fitting goals (3)) and the preconditioned inversion (fitting goals (5)). I am concerned with the behavior in early iterations, so I just ran 6 iterations of each. These results are in Figure 2.

The regularized result is high frequency, but it has barely begun to fill in the areas affected by the null space. This is exactly the behavior we expect at early iterations in a regularized inversion. The preconditioned result has completely filled in the areas affected by the null

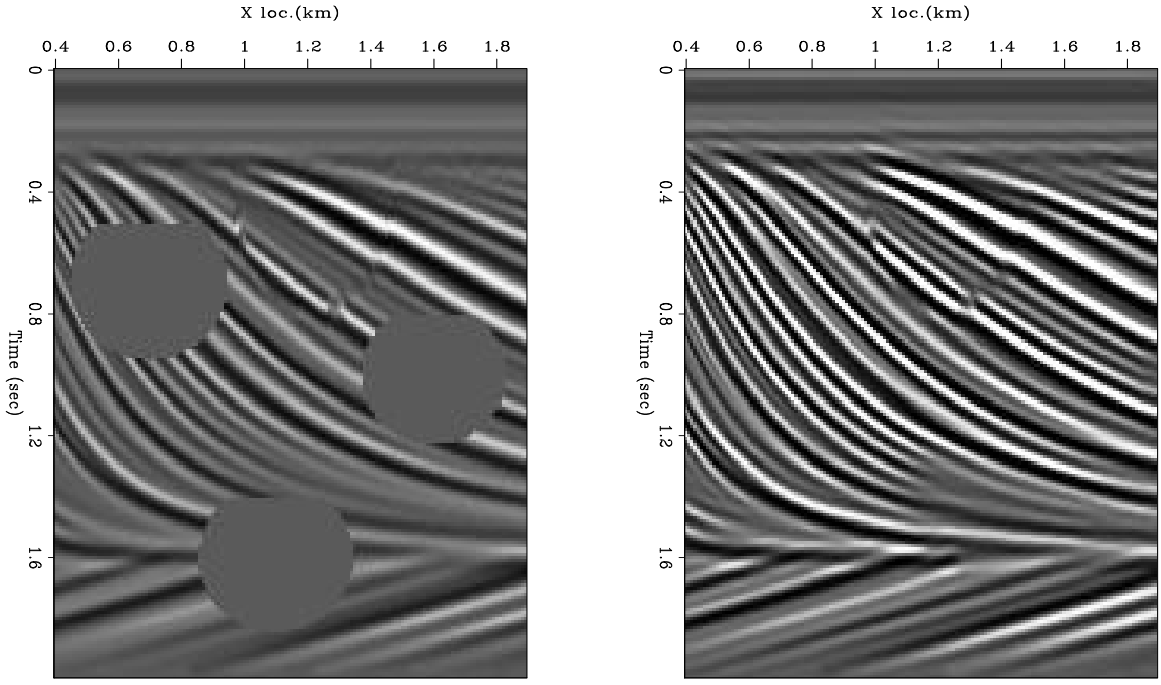


Figure 1: Left panel is the data weighted by the masking operator used for the inversion problems, right panel is the ideal model we get when the masking operator is replaced with an identity operator. marie1-datmod [ER,M]

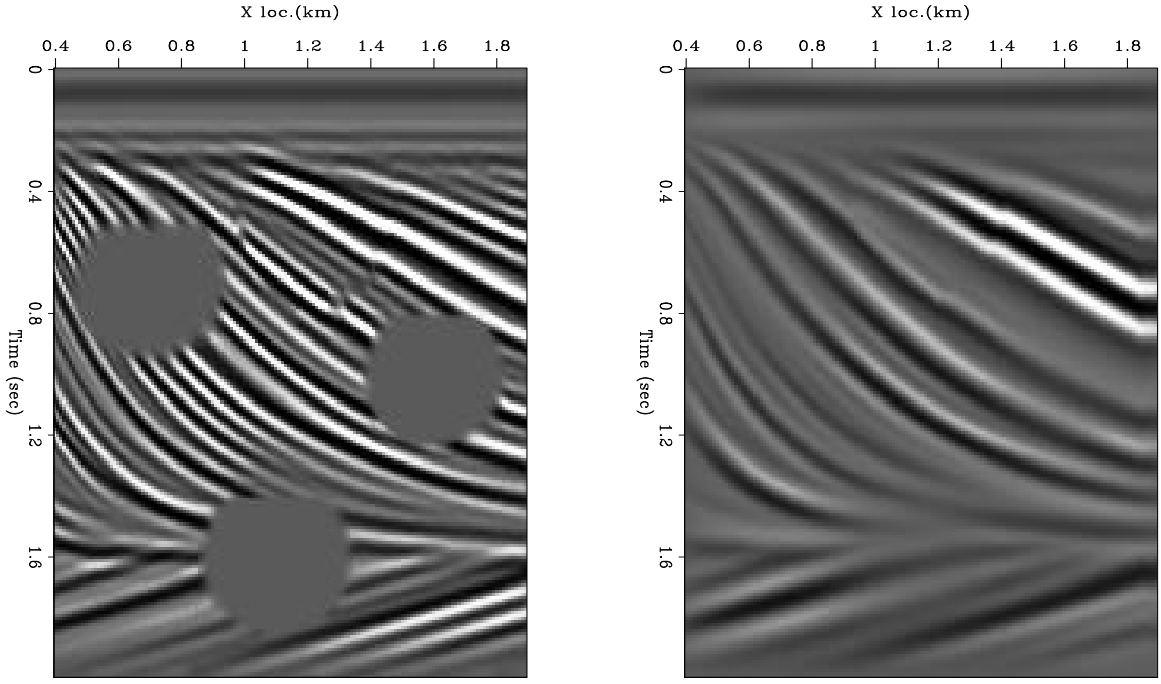


Figure 2: Left panel is the result of 6 iterations of just regularized inversion, right panel is the result of just 6 iterations of preconditioned inversion. marie1-regprec [ER,M]

space, meaning that it has defined solutions at every point (although not the ideal solution), but it is very low frequency. Once again, this is expected and has been seen in earlier works with imaging operators (Prucha et al., 2000; Prucha and Biondi, 2000, 2002).

The previous example helps to demonstrate two important points made by Claerbout (1999). First, both regularized inversion and preconditioned inversion take a great many iterations to converge. While this is not a problem for a toy problem like the one presented in this paper, it is impossible for a real geophysical problem like imaging in complex areas. The operators used in such a problem are infinitely more complex than those used in this simple interpolation problem, so it is vital that we minimize the number of iterations needed (Biondi and Vlad, 2001). Secondly, when we limit ourselves to a small number of iterations, we encounter several problems with both regularization and preconditioning. These problems include:

- A regularized inversion using a small roughening operator will not fill the null space.
- The result of a preconditioned inversion will not contain high frequencies.

Clearly, in order to obtain a high frequency result with defined solutions at every point in a small number of iterations, we need some combination of the regularized and preconditioned inversions. I chose to run a small number of preconditioned iterations then use that result as an initial model for a small number of regularized iterations. I chose to test two different combinations, one with 3 iterations of preconditioned inversion and 3 iterations of regularized inversion and one with 5 iterations of preconditioned inversion and 1 iteration of regularized inversion. These results are in Figure 3.

Both of the CIPR results contain higher frequencies than the purely preconditioned result (right panel Figure 2) and fill the areas affected by the null space better than the purely regularized result (left panel Figure 2). Determining which CIPR result is "better" is fairly subjective, but I chose to compare them by looking at their frequency spectrums. This can be seen in Figure 4. The frequencies shown in this figure are the average over all of the traces.

Figure 4 shows the frequency spectra of the results in Figure 2 and Figure 3 along with the frequency spectrum of the "ideal" model in Figure 1. As expected, the frequency content of the regularized inversion is close to that of the ideal model and the frequency content of the preconditioned inversion is much lower than the ideal model. It is more interesting to compare the frequency contents of the two different CIPR results. This shows us that the inversion using 3 iterations of preconditioning with 3 iterations of regularization has a frequency content closer to the ideal model than that of the inversion using 5 preconditioned iterations and 1 regularized iteration. This is particularly interesting because it indicates that both preconditioning and regularization are important to get the most improvement.

In this paper, I will consider the CIPR result using 3 iterations of preconditioned inversion and 3 iterations of regularized inversion to be my "best" result. Given this result, I felt it would be instructional to see how many iterations of just preconditioned inversion (fitting goals (5)) it would take to get an equivalent frequency content. It took 30 iterations of preconditioned inversion to get the same frequency content as the "best" result. The frequency content of the

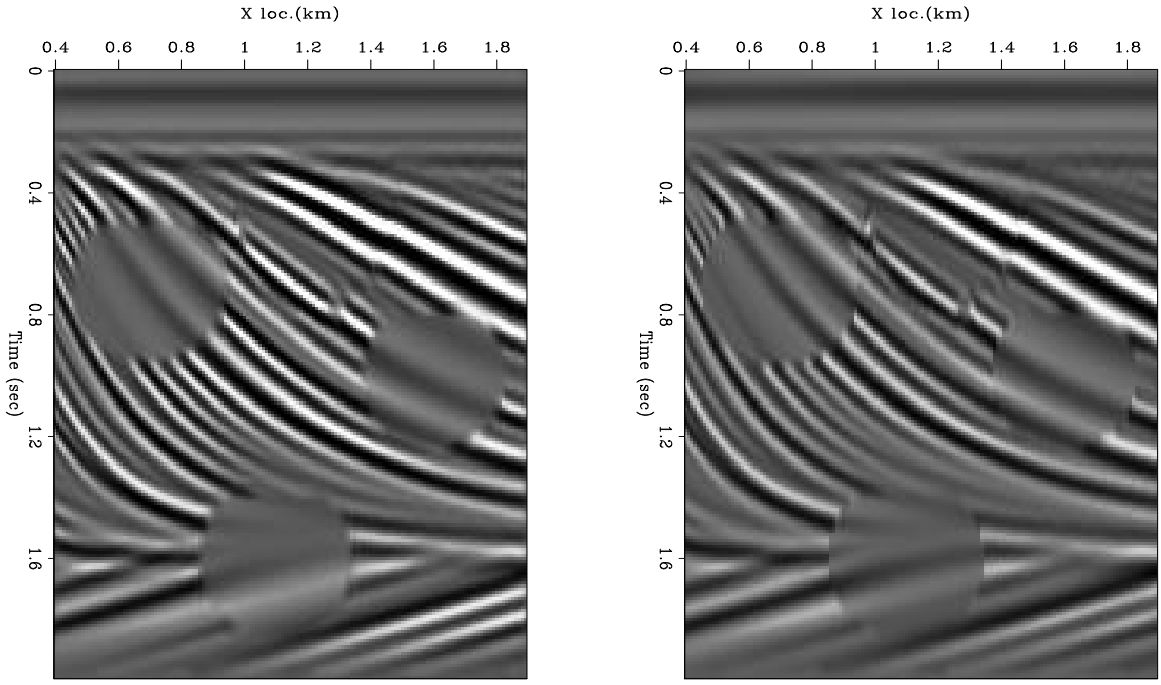


Figure 3: Left panel is the result of 3 iterations of preconditioned inversion followed by 3 iterations of regularized inversion, right panel is the result of 5 iterations of preconditioned inversion followed by only 1 iteration of regularized inversion. marie1-precreg [ER,M]

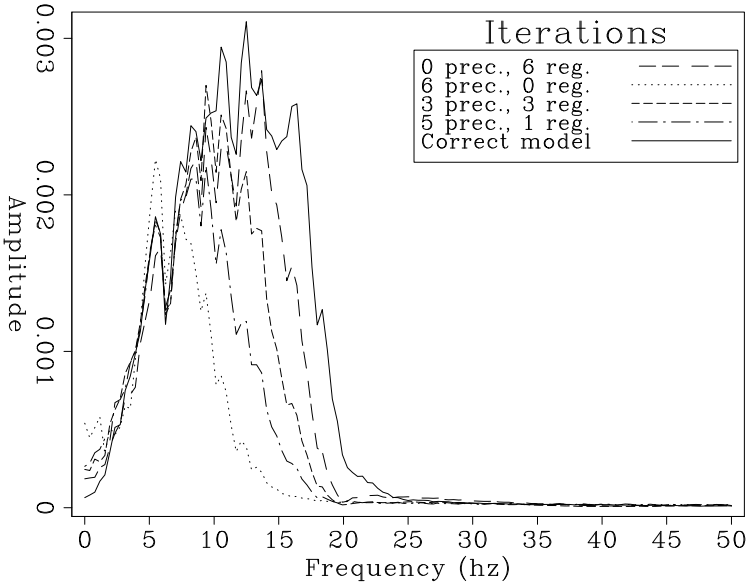


Figure 4: Comparison of the frequency content of the resulting models seen in Figures 2 and 3 along with the frequency content of the correct model (right panel of Figure 1). marie1-spectrum [ER]

result can be seen in Figure 5. One again, the frequencies shown here are the average over all of the traces.

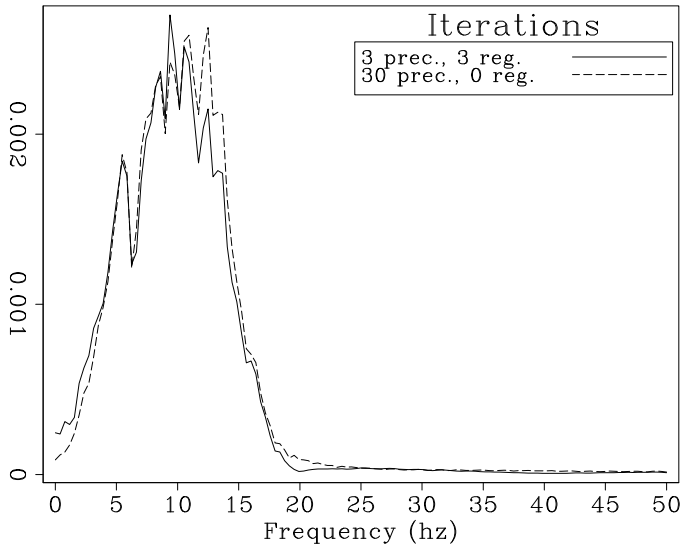


Figure 5: Comparison of the frequency content of the results of 3 iterations of preconditioned inversion with 3 iterations of regularized inversion and 30 iterations of just preconditioned inversion. marie1-speccomp [ER]

Figure 6 displays the models resulting from the "best" solution and the solution using 30 iterations of preconditioned inversion. The model resulting from 30 iterations has done a better job of filling the areas affected by the null space, as we would expect for an inversion process that used 5 times as many iterations. I have also included a model that has filled the areas affected by the null space equally well as that used only regularized inversion (fitting goals (3)). This result took 50 iterations.

CONCLUSIONS

The simple experiment conducted in this paper has compared two familiar inversion schemes, preconditioned and regularized, with a new combined inversion scheme that uses the result of a small number of preconditioned iterations as an initial model for a small number of regularized iterations (CIPR). I used a simple interpolation problem to test CIPR's ability to reduce the number of iterations needed to get an "improved" model. This "improved" model has a solution defined at every point and has a frequency content close to that of an "ideal" model. I have shown that to obtain an "improved" model, CIPR takes far fewer iterations than either of the other schemes. This makes CIPR an interesting option for many types of seismic inversion problems.

FUTURE PLANS

The simple problem presented in this paper will allow me to more thoroughly understand CIPR. One issue I plan to examine is its effect on amplitudes. An extension of this issue is the possibility of using the model normalization described by Rickett (2001a,b) to normalize

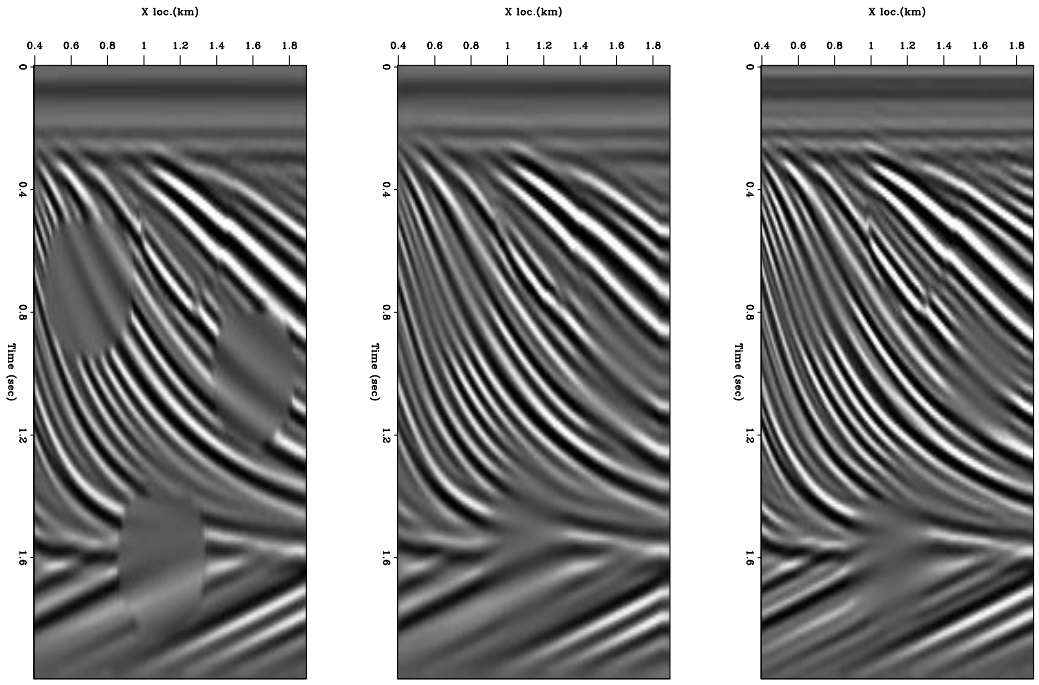


Figure 6: Comparison of the models resulting from 3 iterations of preconditioned inversion with 3 iterations of regularized inversion (left panel), 30 iterations of preconditioned inversion (center panel), and 50 iterations of regularized inversion (right panel). marie1-compits [ER,M]

the result of the preconditioned iterations before it is sent to the regularized inversion as an initial model. Another issue is the possibility of applying a mask that will only allow the preconditioning and regularization to occur within a specified area. Also, I plan to find some way to determine the optimal ratio of preconditioned iterations to regularized iterations. This may be related to another concern I intend to scrutinize, which is the effect of CIPR on the final residuals.

REFERENCES

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

Claerbout, J. F., 1991, Design of inverse Kirchhoff-style filters by LSCG: SEP–**72**, 33–38.

Claerbout, J. F., 1995, Basic Earth Imaging: Stanford Exploration Project, `http://sepwww.stanford.edu/sep/prof/`.

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

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

Fomel, S., 1997, On model-space and data-space regularization: A tutorial: SEP–**94**, 141–164.

Harlan, W. S., 1986, Signal-noise separation and seismic inversion: SEP–**47**.

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

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.

Rickett, J., 2001a, Model-space vs data-space normalization for finite-frequency depth migration: SEP–**108**, 81–90.

Rickett, J., 2001b, Spectral factorization of wavefields and wave operators: Ph.D. thesis, Stanford University.

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

Tikhonov, A. N., and Arsenin, V. Y., 1977, Solution of ill-posed problems: John Wiley and Sons.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 161–171

Short Note

Conjugate gradient total least-squares in geophysical optimization problems

Morgan Brown[1]

INTRODUCTION

Total least-squares (TLS) optimization is a methodology to solve least-squares optimization problems when the modeling operator has errors. In standard least-squares optimization, errors are assumed to be concentrated in the data only.

Golub and Loan (1980) presented a numerically-stable TLS algorithm which utilizes the singular value decomposition (SVD). Subsequent refinements to the method predominantly use SVD, and much of the current literature emphasizes stabilization of the inverse and implicit model regularization by SVD truncation (Fierro et al., 1997). Because it is numerically intensive, however, the SVD generally proves unrealistic for use in large-scale problems, which are the rule in exploration geophysics.

The TLS problem can be cast as an extremal eigenvalue/eigenvector estimation problem. Chen et al. (1986) present a conjugate gradient (CG) scheme to compute the minimum eigenvalue/eigenvector of a linear system. Zhu et al. (1997) extend Chen et al.'s algorithm to solve the TLS problem, in the context of optical tomography.

I begin with a short theoretical overview of the TLS problem. I implement the CG method described by Chen et al. (1986), adapted for the TLS problem in a similar fashion as the work of Zhu et al. (1997). I test the algorithm on two familiar geophysical problems: least-squares deconvolution of a 1-D signal, and velocity scan inversion with the hyperbolic Radon transform. Liu and Sacchi (2002) tested an SVD-based, regularized TLS approach on velocity scan inversion using the parabolic Radon transform.

TLS OVERVIEW

Golub and Loan (1980) phrased the TLS problem as follows. Given a forward modeling operator $\mathbf{L}$ and measured data $\mathbf{d}$, assume that both are contaminated with white noise of uniform

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

variance; matrix **N** and vector **n**, respectively. Then the TLS solution is obtained by minimizing the Frobenius matrix norm of the augmented noise matrix:

$$\min \|[\mathbf{N}\ \ \mathbf{n}]\|_{\mathbf{F}}, \tag{1}$$

subject to the constraint that the solution is in the nullspace of the combined augmented noise and input operators:

$$([\mathbf{L}\ \ \mathbf{d}]+[\mathbf{N}\ \ \mathbf{n}])\begin{bmatrix} \mathbf{m} \\ -1 \end{bmatrix} = \mathbf{0}. \tag{2}$$

To solve the system of equations (1) and (2), Golub and Loan (1980) introduced a technique based on the Singular Value Decomposition (SVD). Although mathematically elegant, SVD-based approaches are generally unrealistic for the large-scale problems that are the norm in exploration geophysics.

Equivalence with Rayleigh Quotient Minimization

Golub (1973) showed that the constrained minimization problem of equations (1) and (2) is equivalent to minimization of the so-called Rayleigh Quotient. If we define the vector $\mathbf{q} = [\mathbf{m}\ \ -1]^T$ and $\mathbf{A} = [\mathbf{L}\ \ \mathbf{d}]$, the Rayleigh Quotient takes the following form:

$$\min\ \ F(\mathbf{q}) = \left| \frac{\mathbf{q}\mathbf{A}^T\mathbf{A}\mathbf{q}}{\mathbf{q}^T\mathbf{q}} \right|_2. \tag{3}$$

After the minimization of equation (3), the resultant vector **q** is the eigenvector associated with the smallest eigenvalue of $\mathbf{A}^T\mathbf{A}$.

Conjugate Gradient Method for TLS

The Rayleigh Quotient can be minimized by iterative techniques. Zhu et al. (1997) introduced a method based on conjugate gradients (CG) to solve the TLS problem which was adapted from the earlier work of Chen et al. (1986). I implemented this CG-based algorithm and present pseudocode in Appendix A.

Theory guarantees that the CG method converges in n steps, where n is the size of the model vector. However, in practical situations with real seismic data, a "useful" model may appear after relatively few ($<< n$) CG iterations. How useful the model and how few the iterations depends on the problem. Nonetheless, in practice, the computational cost and memory requirements are nearly always much less with CG than with SVD.

Relation of TLS to Damped Least-squares (DLS)

The TLS solution is closely related to the classic damped least squares (DLS) solution, where the damping factor, σ^2, is the smallest nonzero singular value of the augmented matrix [**L** **d**]:

$$\mathbf{m}_{DLS} = \left(\mathbf{L}^T\mathbf{L} + \sigma^2\mathbf{I}\right)^{-1}\mathbf{L}^T\mathbf{d}. \tag{4}$$

The TLS solution can be rewritten (Golub and Loan, 1980; Björck, 1996) as follows.

$$\mathbf{m}_{TLS} = \left(\mathbf{L}^T\mathbf{L} - \sigma^2\mathbf{I}\right)^{-1}\mathbf{L}^T\mathbf{d}. \tag{5}$$

The only difference between equations (4) and (5) is the negative sign on the damping term. Thus the TLS problem is considered a "deregularization" of the standard LS problem, and is guaranteed to be worse conditioned, since $\mathbf{L}^T\mathbf{L}$ is positive-semidefinite at worst (Björck, 1996).

LEAST-SQUARES DECONVOLUTION TESTS

I constructed a simple, yet relevant synthetic test case for the TLS algorithm: deconvolution. The known model is a sequence of spikes of random amplitude and placement. To create data, the known model was convolved with a Ricker wavelet. Gaussian-distributed noise with a variance of 1 was added to the data, and also to the filter used in the deconvolution.

Figures 1-3 compare the standard least-squares (LS), the TLS, and DLS solutions to the problem. The LS solution is undoubtedly poor. In the "quiet" zones of the model, where the known model is zero-valued, the estimated LS model has almost as much energy as where the spikes are. Still, the modeled data appears to fit the input data quite well.

The TLS and DLS solutions appear somewhat similar. Both approaches seem to suppress unwanted noise in the estimated model in the quiet regions. However, the TLS model seems to have better resolution of the true spikes. Also, the TLS method's residual error appears better balanced than the DLS's. Both TLS and DLS have higher residual error energy than the LS solution.

HYPERBOLIC RADON TRANSFORM TESTS

I tested the proposed TLS algorithm on a popular SEP inversion application, the Hyperbolic Radon Transform (HRT) (Nichols, 1994; Lumley et al., 1995; Guitton, 2000b). Figures 4 and 5 compare the results of the TLS, LS, and DLS methods, for 10 and 150 CG iterations, respectively.

The results of the HRT tests are inconclusive. After 10 iterations, the results from the three methods are almost indistinguishable. After 50, the DLS model looks "best," i.e., most interpretable by a human for picking velocities. However, the TLS residual error is the whitest, the best balanced, and contains no correlated energy–the very criteria which Guitton (2000a) uses to define optimality.

CONCLUSIONS AND DISCUSSION

I have introduced total least-squares (TLS) optimization as a possible alternative to "standard" least-squares approaches. TLS approaches incorporate errors in both the data and in the mod-

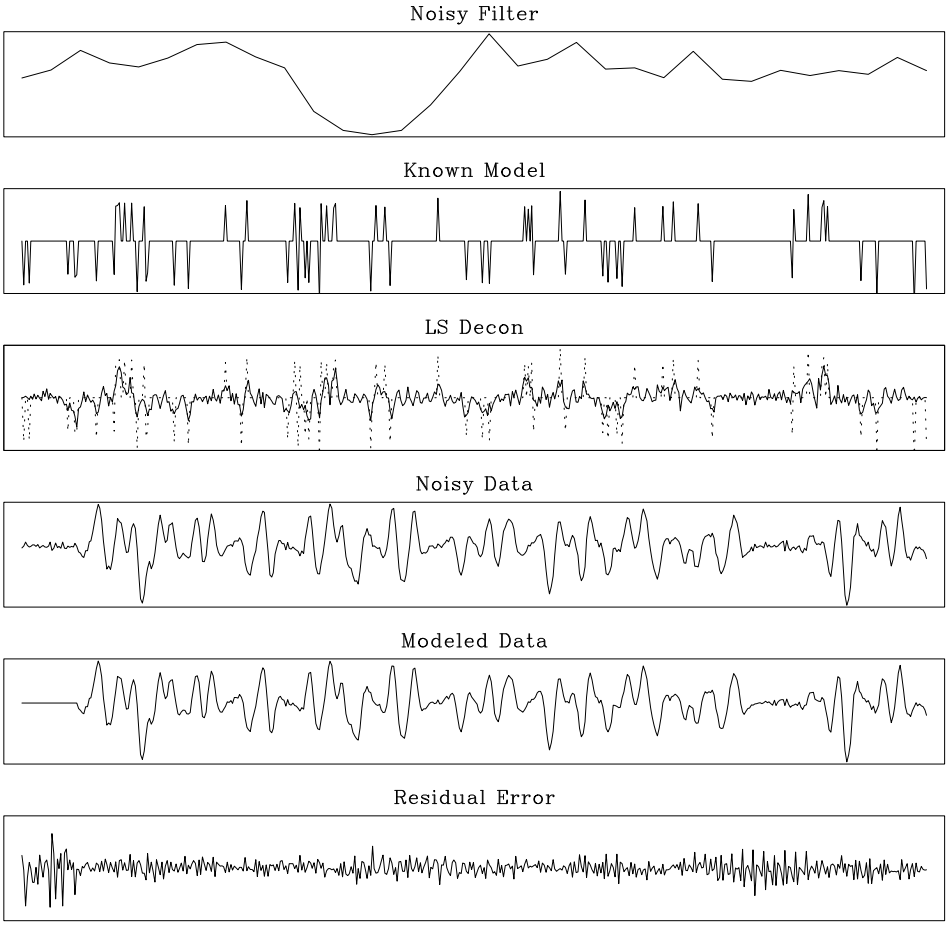


Figure 1: Top to bottom: 1) Known filter plus noise, 2) Known model, 3) Estimated standard least-squares model overlaying known model, 4) Noisy data, 5) Modeled data, 6) Residual error. morgan2-decon.ls.noisy [ER]

eling operator, to produce "more accurate" solutions. I put "more accurate" in quotes because in our real world, frequent appearences of nonempty nullspaces impose some subjectivity on any solution.

I implemented the conjugate gradient TLS solver (TLS-CG) published by Zhu et al. (1997), although in that paper, the authors omit a crucial model normalization step that leads to non-convergence of the algorithm. I present a complete algorithm in Appendix A.

Tests on a synthetic 1-D deconvolution example seem to validate TLS as a tool. In those tests, when ideal noise was added to the filter and data, TLS resolved the true model better than normal least-squares or damped least-squares. Tests using the hyperbolic radon transform were inconclusive; no efforts were made to understand operator error in this case, and in summary, the TLS result looks somewhere in between LS and DLS.

Will TLS be a useful tool in geophysics? My suspicion is that TLS makes only a second order improvement in the quest to account for uncertainty in geophysical inverse problems.

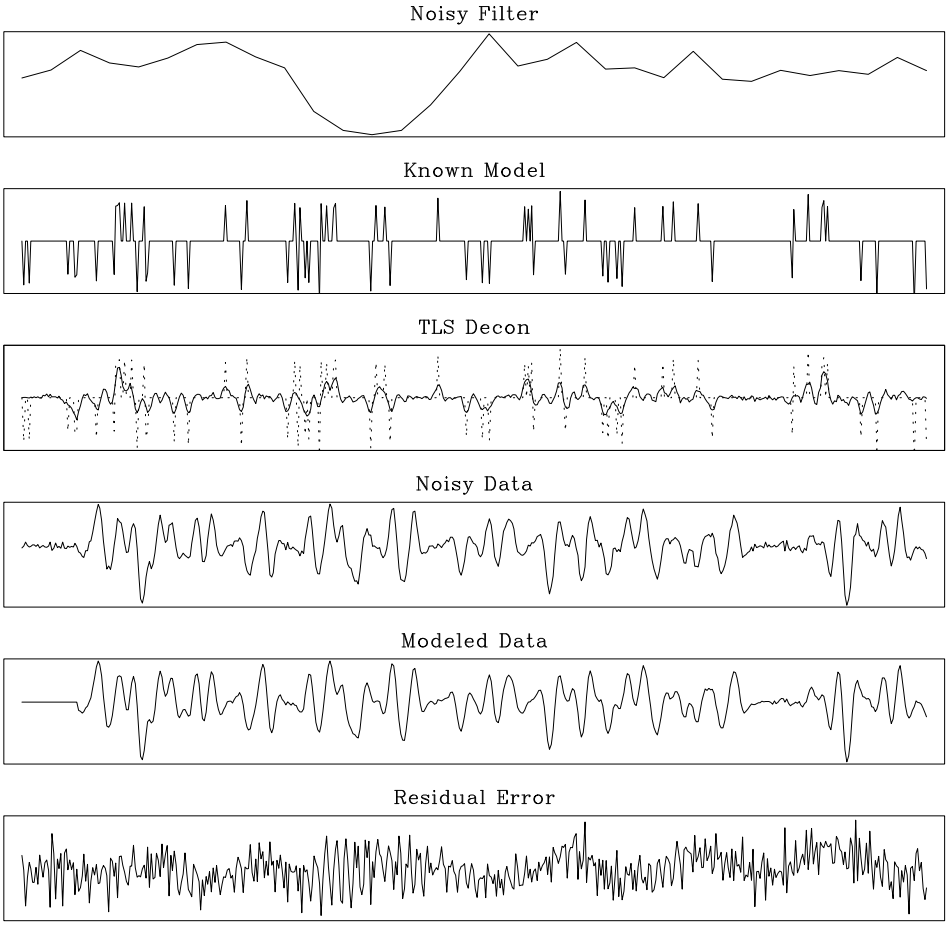


Figure 2: Top to bottom: 1) Known filter plus noise, 2) Known model, 3) Estimated total least-squares model overlaying known model, 4) Noisy data, 5) Modeled data, 6) Residual error. morgan2-decon.tls.noisy [ER]

More interesting are efforts to perturb the nullspace of inverse problems to infer model statistics (Clapp, 2002; Chen and Clapp, 2002).

Discussion: Error Distribution

Recall that in the earlier TLS formulation, the noise which contaminates *both* the operator and data is assumed to be white, with uniform variance. In practice, both the operator and data noise are likely to be correlated, with nonuniform variance. Björck (1996) notes that an appropriate change of variables can restore the validity of the assumptions. He defines a square matrix $\mathbf{D}$ which is applied, somewhat surprisingly, to the "data matrix" of equation (2).

$$(\mathbf{D}[\mathbf{L}\ \ \mathbf{d}] + [\mathbf{N}\ \ \mathbf{n}]) \begin{bmatrix} \mathbf{m} \\ -1 \end{bmatrix} = \mathbf{0}. \tag{6}$$

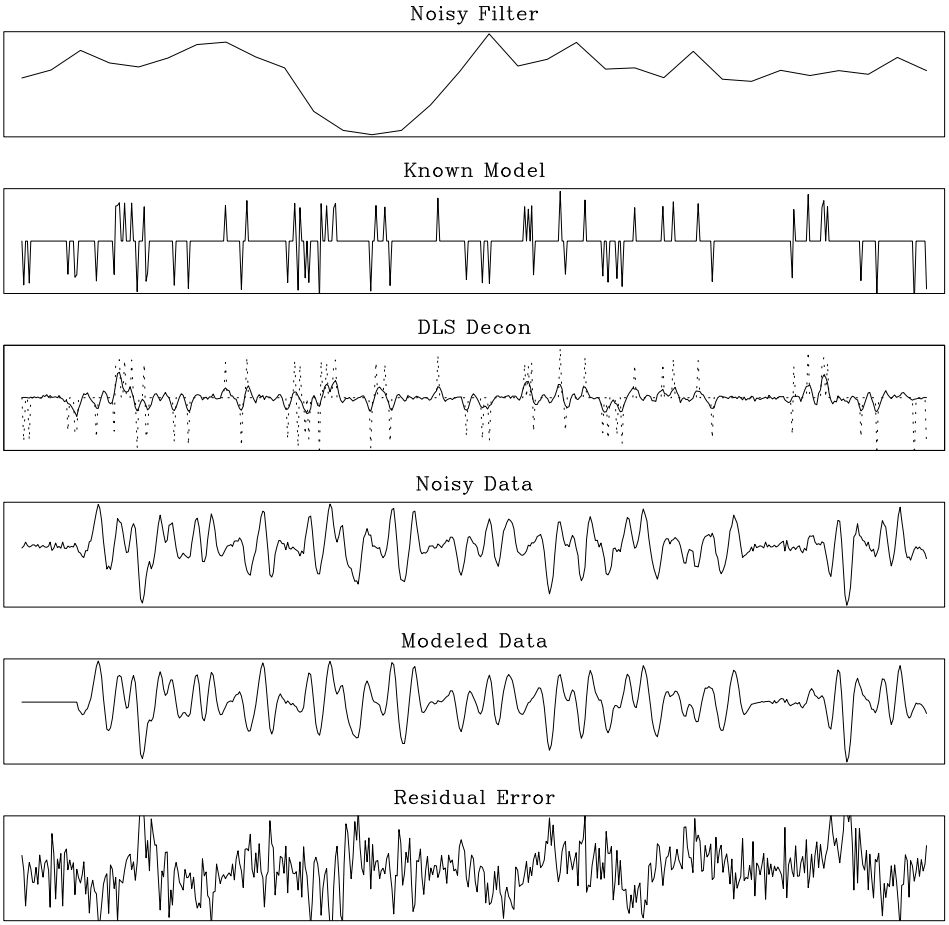


Figure 3: Top to bottom: 1) Known filter plus noise, 2) Known model, 3) Estimated damped least-squares model overlaying known model, 4) Noisy data, 5) Modeled data, 6) Residual error. morgan2-decon.dls.noisy [ER]

While it may seem intuitive to scale the noise, rather than the data, if the operator is diagonal (as it is in the fairytale world of uncorrelated noise), the inverse is trivial. Even if the noise is correlated, at SEP, we have considerable experience with the design of invertible decorrelation and balancing operators.

Are the restrictions (white, balanced) on the noise crippling? Zhu et al. (1997) claim that in scattering tomography experiments, correlated noise does not unduly harm the TLS result, and also that the TLS result in this case is still better than the normal LS result.

ACKNOWLEDGEMENT

At the 2002 SEP sponsor meeting, Peter Harris of CGG suggested I look into TLS.

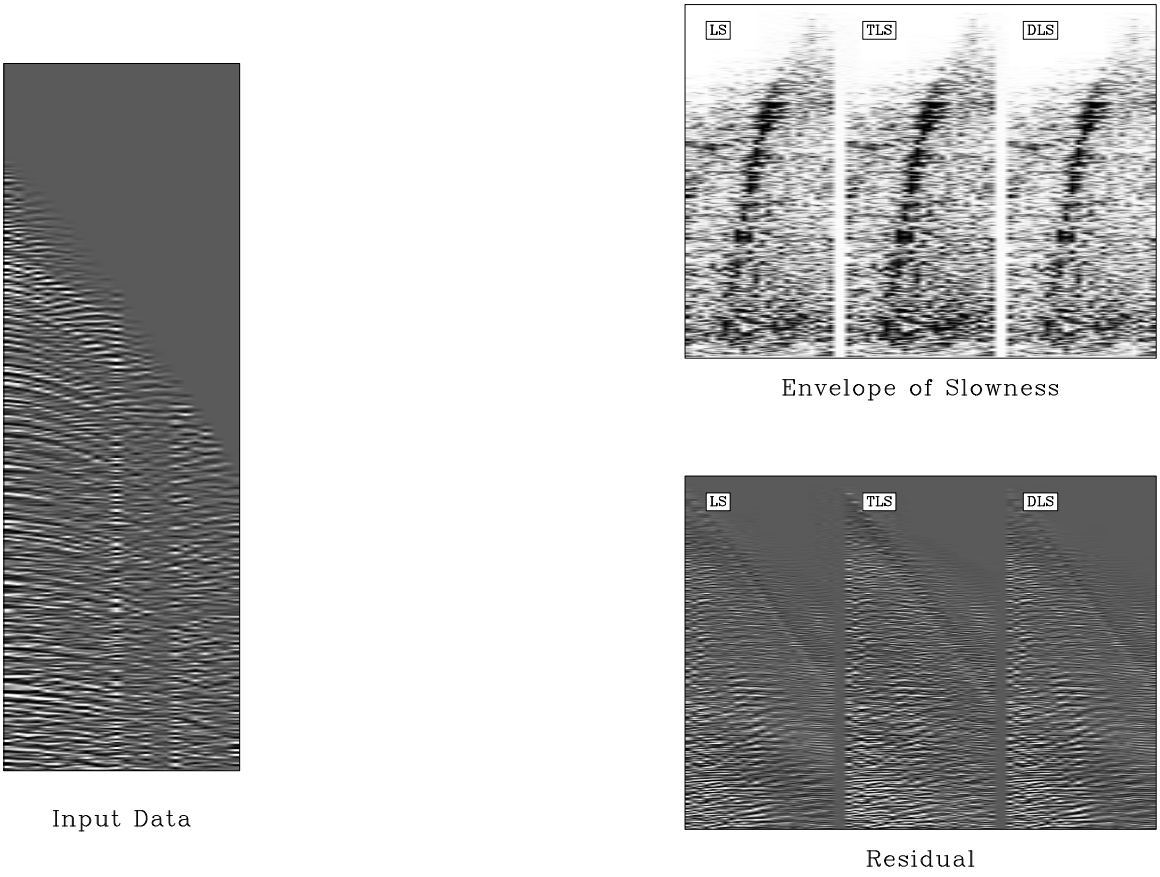


Figure 4: Left panel: Input data. Right-top: Envelope of estimated slowness model for LS, TLS, and DLS methods after 10 iterations. Right-bottom: Residual error for LS, TLS, and DLS solutions. morgan2-hrtcomp.10 [ER]

REFERENCES

Björck, A., 1996, Numerical methods for least squares problems: Society for Industrial and Applied Mathematics.

Chen, W., and Clapp, R., 2002, Exploring the relationship between uncertainty of ava attributes and rock information: SEP–**112**, 259–268.

Chen, H., Sarkar, T., Dianat, S., and Brulé, J., 1986, Adaptive spectral estimation by the conjugate gradient method: IEEE Transactions on Acoustics, Speech, and Signal Processing, **ASSP-34**, no. 2, 272–284.

Clapp, R. G., 2002, Effect of migration velocity uncertainty on amplitude information: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Fierro, R., Golub, G., Hansen, P., and O'Leary, D., 1997, Regularization by truncated total least squares: SIAM J. Sci. Comput., **18**, no. 4, 1223–1241.

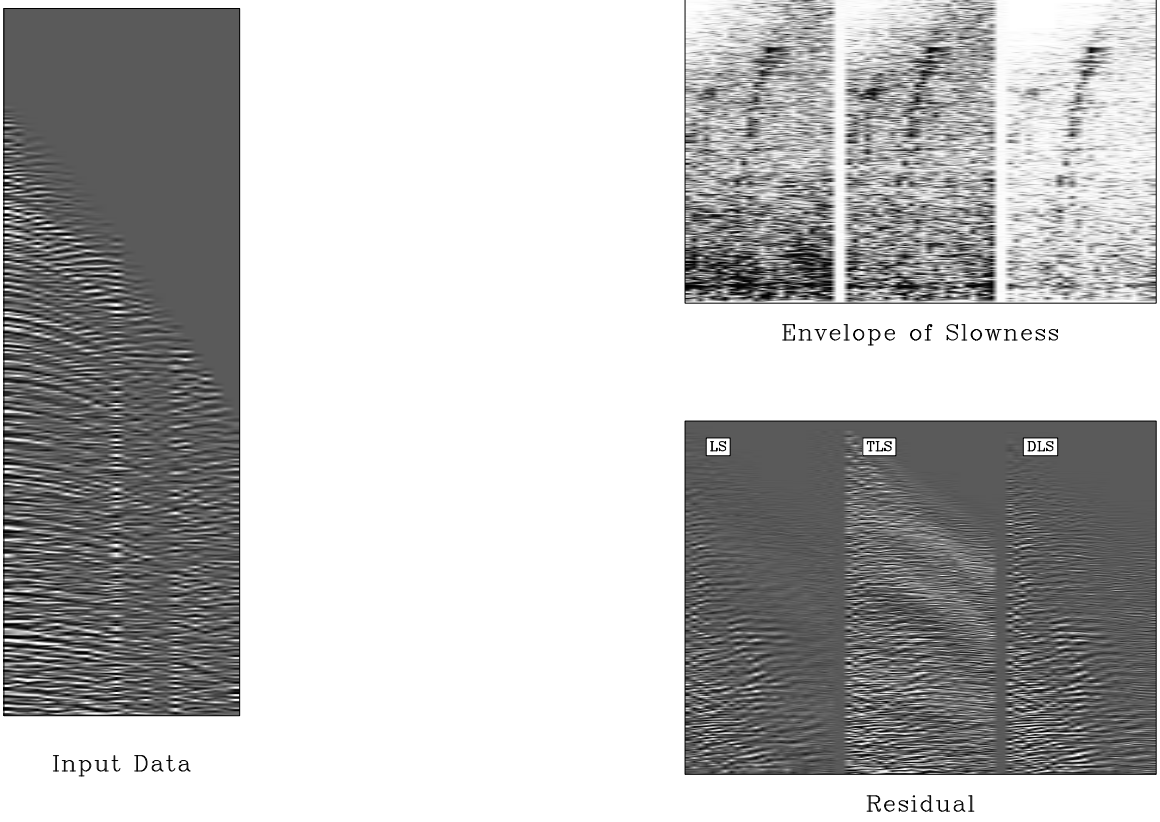


Figure 5: Left panel: Input data. Right-top: Envelope of estimated slowness model for LS, TLS, and DLS methods after 50 iterations. Right-bottom: Residual error for LS, TLS, and DLS solutions. morgan2-hrtcomp.50 [ER]

Golub, G., and Loan, C. V., 1980, An analysis of the total least squares problem: SIAM J. Numer. Anal., **15**, no. 17, 883–893.

Golub, G., 1973, Some modified matrix eigenvalue problems: SIAM (Soc. Ind. Appl. Math.) Rev., **15**, no. 15, 318–334.

Guitton, A., 2000a, Coherent noise attenuation using Inverse Problems and Prediction Error Filters: SEP–**105**, 27–48.

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

Liu, Y., and Sacchi, M., 2002, Regularization of inverse problems using total least squares: Ann. Mtg, Can. Soc. Expl. Geophys., Expanded Abstracts, pages 1–3.

Lumley, D. E., Nichols, D., Ecker, C., Rekdal, T., and Berlioux, A., 1995, Amplitude-preserved processing and analysis of the Mobil AVO data set: SEP–**84**, 125–152.

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

Zhu, W., Wang, Y., Yao, Y., Chang, J., Graber, H., and Barbour, R., 1997, Iterative total least-squares image reconstruction algorithm for optical tomography by the conjugate gradient method: J. Opt. Soc. Am. A, **14**, no. 4, 799–807.

APPENDIX A: CONJUGATE GRADIENT MINIMIZATION OF THE RAYLEIGH QUOTIENT

Recall that $\mathbf{q} = [\mathbf{m} \quad -1]^T$, where $\mathbf{m}$ is the "usual" model (i.e., **Lm=d**). $\mathbf{q}_i$ is the estimated model vector at iteration `i`.

$$\mathbf{q}_0 = \frac{\mathbf{q}_0}{\sqrt{\mathbf{q}_0^T \mathbf{q}_0}} \quad \Leftarrow \text{Model} \tag{A-1}$$

$$\lambda_0 = \mathbf{q}_0^T \mathbf{A}^T \mathbf{A} \mathbf{q}_0 \quad \Leftarrow \text{Estimated smallest eigenvalue} \tag{A-2}$$

$$\mathbf{r}_0 = \lambda_0 \mathbf{q}_0 - \mathbf{A}^T \mathbf{A} \mathbf{q}_0 \quad \Leftarrow \text{Residual} \tag{A-3}$$

$$\mathbf{s}_0 = \mathbf{r}_0 \quad \Leftarrow \text{Solution Step} \tag{A-4}$$

$$\text{iterate} \quad \{ \quad (k = 0, n_{iter}) \tag{A-5}$$

$$P_{a,k} = \mathbf{q}_k^T \mathbf{A}^T \mathbf{A} \mathbf{s}_k \tag{A-6}$$

$$P_{b,k} = \mathbf{s}_k^T \mathbf{A}^T \mathbf{A} \mathbf{s}_k \tag{A-7}$$

$$P_{c,k} = \mathbf{s}_k^T \mathbf{q}_k \tag{A-8}$$

$$P_{d,k} = \mathbf{s}_k^T \mathbf{s}_k \tag{A-9}$$

$$b = P_{b,k} - \lambda_k P_{d,k} \tag{A-10}$$

$$c = P_{a,k} - \lambda_k P_{c,k} \tag{A-11}$$

$$d = P_{b,k} P_{c,k} - P_{a,k} P_{d,k} \tag{A-12}$$

$$\alpha_k = \frac{-b + \sqrt{b^2 - 4dc}}{2d} \tag{A-13}$$

$$\mathbf{q}_{k+1} = \mathbf{q}_k + \alpha_k \mathbf{s}_k \tag{A-14}$$

$$\mathbf{q}_{k+1} = \frac{\mathbf{q}_{k+1}}{\sqrt{\mathbf{q}_{k+1}^T \mathbf{q}_{k+1}}} \tag{A-15}$$

$$\lambda_{k+1} = \mathbf{q}_{k+1}^T \mathbf{A}^T \mathbf{A} \mathbf{q}_{k+1} \tag{A-16}$$

$$\mathbf{r}_{k+1} = \lambda_{k+1} \mathbf{q}_{k+1} - \mathbf{A}^T \mathbf{A} \mathbf{q}_{k+1} \tag{A-17}$$

$$\beta_k = -\frac{\mathbf{s}_k \mathbf{A}^T \mathbf{A} \mathbf{r}_{k+1}}{\mathbf{s}_k \mathbf{A}^T \mathbf{A} \mathbf{s}_k} \tag{A-18}$$

$$\mathbf{s}_{k+1} = \mathbf{r}_{k+1} + \beta_k \mathbf{s}_k \tag{A-19}$$

$$\} \tag{A-20}$$

$$\mathbf{q}_{n_{iter}} = \mathbf{q}_{n_{iter}} / \left(-\mathbf{q}_{n_{iter}} [m+1] \right) \tag{A-21}$$

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 171–181

Nonlinear pairwise alignment of seismic traces

Christopher L. Liner and Robert G. Clapp[1]

ABSTRACT

Alignment of seismic traces is a recurring need in seismic processing and interpretation. For global alignment via static shift there are robust tools available, including cross correlation. However, another kind of alignment problem arises in applications as diverse as associating synthetic seismograms to field data, harmonizing P-wave and mode converted data, final multilevel flattening of common image gathers, and so on. These cases require combinations of trace compression, extension, and shift - all of which are time variant. The difficulty is to find a mapping between the traces which is in some sense optimum. This problem is solved here using a modified form of the Needleman-Wunsch algorithm, which was originally developed for amino acid sequence alignment in proteins. Applied to seismic traces, this global optimization algorithm provides a nonlinear mapping of one seismic trace onto another. The method extends to alignment of any number of traces since that problem can be broken down into a cascade of pairwise alignments. The Needleman-Wunch algorithm is discussed, extended to the seismic case, and applied to field data. The results show a promising new tool for nonlinear alignment or flattening of seismic data.

INTRODUCTION

This paper is concerned with the process of aligning two seismic traces and, by repeated application, any number of seismic traces. At first appearance this is a trivial problem. One need only compute the cross correlation of the two traces and, from the peak of this function, the optimum alignment is known.

But this is merely the best alignment via static shift. In a complex data set, such as a CMP gather, the human eye can associate events that a global correlation alignment will not honor. What the eye is able to do is a time-variant, nonlinear association of events. To carry out this alignment process requires the determination and application of a nonlinear mapping between the trace samples – a combination of compression, stretching, and translation, all varying with time.

If it were possible to do this kind of alignment, what would be the use of it? In a sense, it is already done everyday in seismic processing. A collection of traces are analyzed for a set

[1]**email:** cll@utulsa.edu, Department of Geosciences , University of Tulsa; bob@sep.stanford.edu

of coefficients which drive a nonlinear stretch to make all events flat at all times. This process is, of course, normal moveout. But NMO is a model-driven process, with the model being the NMO equation. A general nonlinear trace alignment algorithm would make it possible to flatten all events in a CMP gather with no knowledge of the NMO equation. We are not advocating such a procedure, but making a point. There is value in using the NMO equation to flatten events, including the fact that it leaves multiples non-flat and therefore removable (at least partially). However, a the general alignment algorithm may be useful as a final flattening procedure for any type of gather (CMP, CIG, CAG, etc.). It could also have application in alignment of synthetic seismograms with field data, associating events on P-P and P-Sv data cubes, etc. In short it would be a useful, general utility.

While the literature on trace interpolation and estimation of missing data is vast, there is very little published work on nonlinear trace alignment. To our knowledge the only published work directly on point with the current study is Martinson et al. (1982) and a derivative paper Martinson and Hopper (1992). In the second work, an iterative, linear inverse approach is used to determine a set of coefficients describing a mapping function which relates features on one trace with those on another. The process is driven by maximizing the correlation or coherence between the modified traces, and used as a trace interpolation technique. This method is similar in spirit to our approach, but owing to the use of linear inverse theory it tends to be expensive, sensitive to the starting model, and does not guarantee a global solution.

Our solution to the pairwise trace alignment problem borrows a concept and algorithm from computational biology and modifies it to the seismic case. The concept is pairwise alignment of amino acid sequences, and the algorithm is due to Needleman and Wunsch (1970).

METHODOLOGY

Needleman-Wunsch algorithm

The Needleman-Wunch (NW) algorithm (Needleman and Wunsch, 1970) is a nonlinear global optimization method that was developed for amino acid sequence alignment in proteins. This was the first of many important alignment techniques which now find application in the Human Genome Project.

Human DNA consists of some 30,000 genes which are in turn composed of 20 amino acids represented by letters of a reduced alphabet (ADCEFGHILKMNPQRSTVWY). The total genome is composed of about 3 billion chemical base pairs, or about $100,000$ per gene. Finding where a particular string of amino acids fits is an optimization problem that aims to find the optimal alignment of the two strings with respect to a defined set of rules and parameter values for comparing different alignments.

The algorithm is an iterative method in which all possible pairs of amino acids (one from each string) are set up in a 2D matrix and alignments are represented as pathways through this array. The optimum alignment is the path (or paths) connecting maximum scoring values. This approach is an example of dynamic programming, which has also been applied to seismic

modeling (Darby and Neidell, 1966) and travel time computation (Schneider et al., 1992).

It is a global optimization process which yields a solution to the problem of pairwise alignment, meaning that we are interested in finding the best fit between only two strings. If alignment of more than two strings is of interest, the problem can be broken down into a cascade of pairwise alignments and thus solved.

In its simplest form, the Needleman-Wunsch algorithm can be summarized by Figure 1. A matrix is formed by placing the two strings, possibly of different length, along the left column and top row. In this step a one is allocated to a cell in the matrix if the letter in each list at this location is the same. Otherwise no entry is made (which is a defacto zero). It is at this stage that the letter-alignment problem becomes purely numerical. In fact, the original string could just as easily consist of integers as letters. The result of this process is the similarity matrix in Figure 1a.

From the similarity matrix a scoring matrix is formed beginning in the lower right corner. The procedure is to add the score value to the maximum value in a row-column pair whose upper left corner is down and to the right of the current working position. Thus in Figure 1b the similarity value 1 is added to the maximum value in the blackened cells (also 1) to give a score of 2. Figure 1c is a later stage of the computation, which continues up and to the left until every cell has been visited and the scoring matrix is complete, Figure 1d. In this simple form, a final score corresponds to how many character matches exist in the optimum alignment.

The final step (traceback) operates by starting at the highest score value (8 in this case) and determining the maximum score path by moving to the right, down, or diagonally down and to the right, Figure 1e. The fact that more than one 8 score alignment exists (Figure 1f) is an expression of non-uniqueness. An important aspect of the solution is that in the process of finding the best global alignment, we also find the best alignments of any sublength.

Details of the algorithm

We now introduce a more flexible form to the Needleman-Wunsch algorithm (Karp, 2000). Let the two input strings, (x, y), given by

$$x = (x_1, x_2, \ldots, x_i, \ldots, x_m) \tag{1}$$
$$y = (y_1, y_2, \ldots, y_j, \ldots, y_n) \tag{2}$$

where (m,n) need not be equal. The subscript i denotes the row direction, and j denotes columns. We write the scoring function $V(i, j)$ as the equation set

$$V(i,j) = max\left[G(i,j), F(i,j), E(i,j)\right] \tag{3}$$
$$G(i,j) = \sigma(x_i, y_j) + V(i+1, j+1) \tag{4}$$
$$F(i,j) = -(p+q) + max\left[V(i+1,j), F(i+1,j)+p\right] \tag{5}$$
$$E(i,j) = -(p+q) + max\left[V(i,j+1), E(i,j+1)+p\right] \tag{6}$$
$$V(i,n+1) = -(p+(m-i+1)\,q) \tag{7}$$
$$V(m+1,j) = -(p+(n-j+1)\,q) \quad , \tag{8}$$

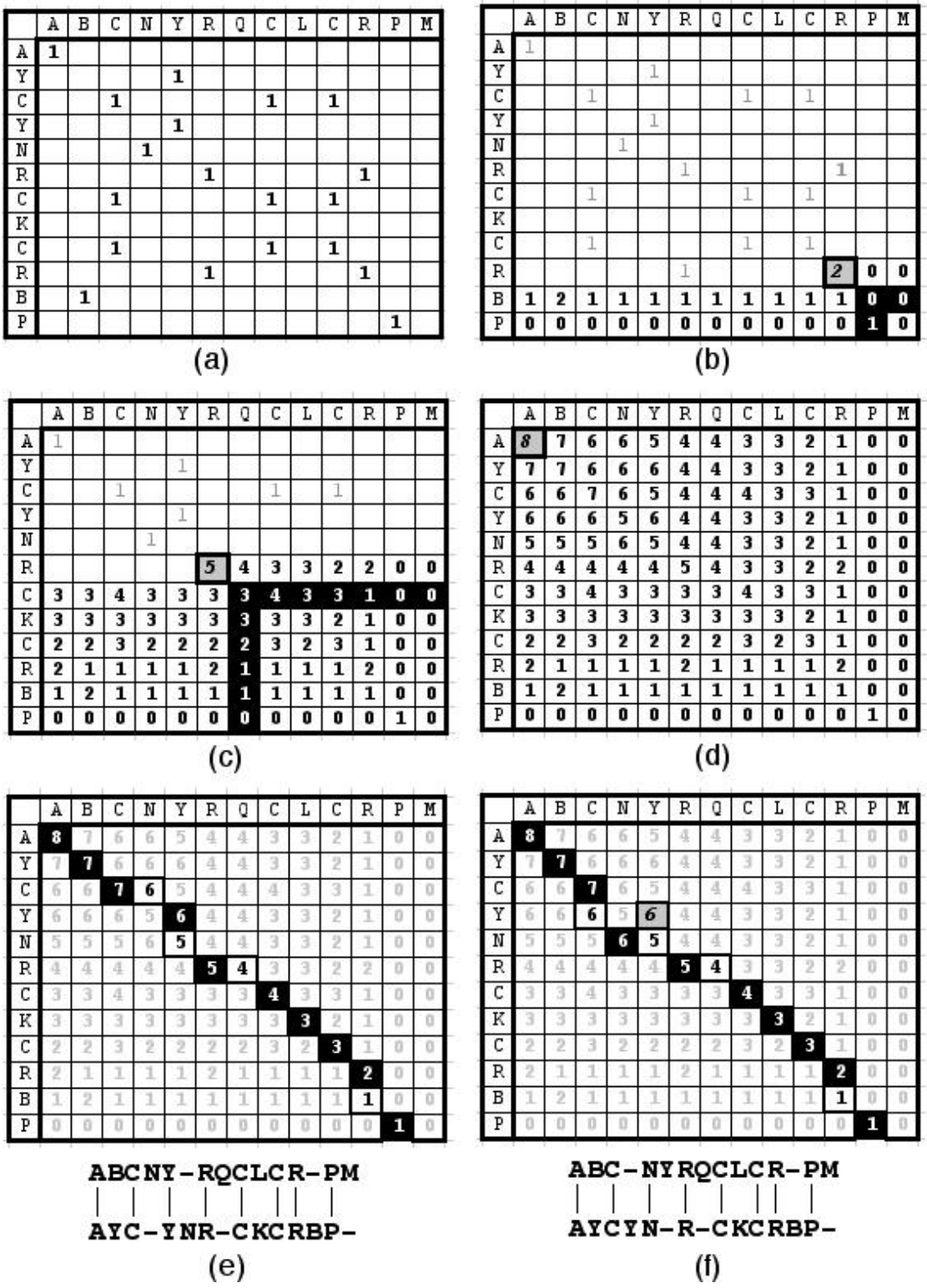


Figure 1: The Needleman-Wunsch algorithm yields the globally optimum alignment between two strings, one along the left of the matrix and the other across the top. (a) Similarity matrix. (b) and (c) partially complete score matrix. (d) Complete score matrix. (e) Traceback route giving globally optimum alignment. (f) Alternate alignment illustrating non-uniqueness. bob2-algo [NR]

where the indices range $i \leq m$ and $j \leq n$. In this form, the function $\sigma(x_i, y_j)$ is the similarity matrix and it is calculated on the fly rather than precomputed. Further it can be customized to reflect different weights associated with matches and mismatches. In biological applications an element on each string either matches or does not, and this fact is represented in the choice of a similarity measure, for example

$$\sigma(x_i, y_j) = \begin{cases} \sigma(a,a) = +1 \\ \sigma(-,a) = 0 \\ \sigma(a,-) = 0 \end{cases} \tag{9}$$

in which a value of 1 is awarded for a match, and all other cases are are equally awarded 0.

In the seismic case we do not expect or need an exact amplitude match, rather it is important to reward small amplitude differences and penalize large ones. We capture this idea in a similarity function as

$$\sigma(x_i, y_j) = c - abs[t_1(x_i) - t_2(y_j)], \tag{10}$$

where c is a constant chosen to keep $\sigma(x_i, y_j) > 0$ and $abs[]$ is the absolute value.

Computational complexity and cost of this algorithm applied to two strings of length n and m is $O(n * m)$.

Modification to the seismic case

On first consideration, the NW algorithm seems ill-suited to the seismic case, primarily because seismic amplitude data is continuous, not discrete. To utilize the machinery of the NW algorithm we consider the histogram of data amplitudes on the trace pair and form a set of bins. That is to say, all of the floating point amplitudes are partitioned into a small number of intervals (20 in the examples given below), and the similarity matrix is formed by the equation 10 operating on the binned amplitudes. An important aspect of the algorithm is that a global optimum alignment function is found independent of the similarity measure that is used. To test feasibility, we used a one point similarity that captures amplitude differences. However we could easily have worked with a twopoint measure to emphasize slope similarity, or three points to match curvature. Extending this idea, one could work with short window cross correlations to fill the similarity matrix similar (Martinson and Hopper, 1992). Clearly, any of these more ambitious similarity measures would increase the cost of the algorithm. In any case, the NW algorithm guarantees a global optimum alignment solution using any similarity matrix as input.

As a final comment, we note that any number of traces that require alignment can be processed as a cascaded series of pairwise problems. Thus there is no loss of generality in discussing just the pairwise problem.

EXAMPLES

To test the methodology we started with a simple synthetic Normal Moveout (NMO) gather with some random noise (Figure 2). We selected two traces some distance apart (the first and tenth trace) and applied the algorithm. The two traces can be seen in the left part of Figure 3. Note the time-variant alignment error.

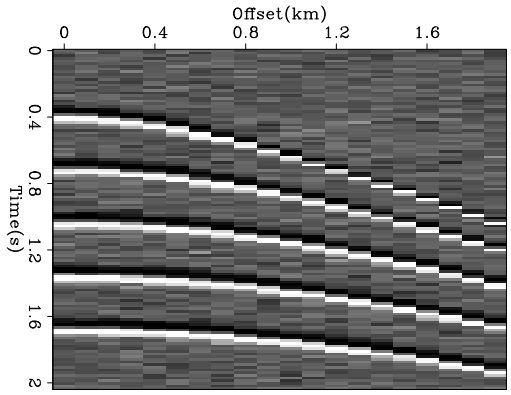


Figure 2: A synthetic shot gather. The first and tenth trace were selected to test the algorithm. bob2-nmo.gather [ER]

Figure 4 shows the similarity (left) and score (right) matrices. The black lines in the similarity matrix represent low scores and correspond to the events in the data, they only disappear (or match) when encountering another wavelet. The score matrix shows exactly what we expect to see. A slightly non-diagonal maximum (except for edge effects at low times corresponding to a lack of coherent events). Figure 3 shows the input (left) and output (right) along with their corresponding differences. The output is much better aligned and the overall differences reduced. The difference trace is a proxy for guaging the quality of alignment, but the goal is not to drive this difference to zero. The algorithm keys on strong events whose alignment may result in sizable differences at other levels. This is a significant departure from Martinson and Hopper (1992) who minimize a difference measure to determine alignment.

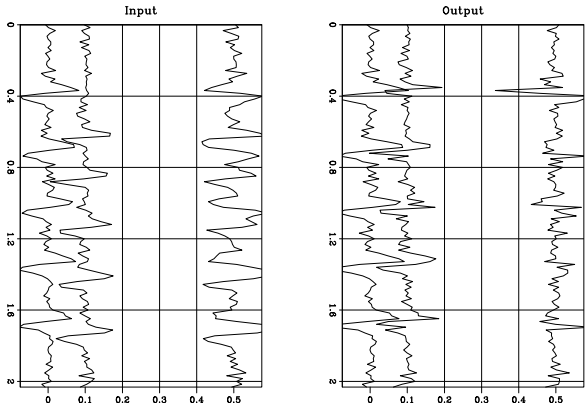


Figure 3: The left plot shows the two input traces and the right plot the traces after alignment. The third trace in each display shows the difference. bob2-nmo.in-out [ER]

For a second test we chose a common reflection point (CRP) gather from a 2-D marine dataset (Figure 5). The gather is an angle gather (Prucha et al., 1999; Sava and Fomel, 2000) after phase-shift plus-interpolation (PSPI) migration (Gazdag and Sguazzero, 1985). Note that

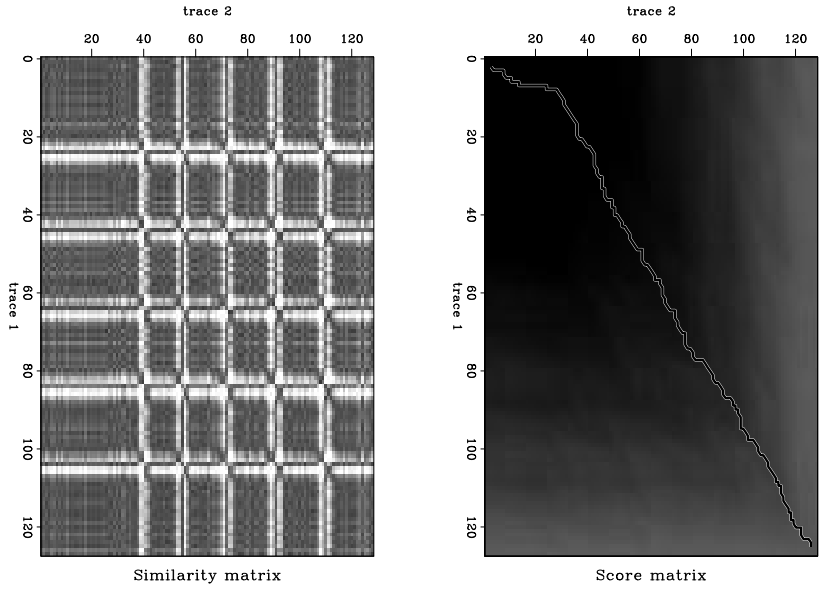


Figure 4: The right plot is the score matrix using first two traces from Figure 2 and the left panel is the similarity matrix. Axes labels refer to time sample numbers (not seconds). bob2-nmo.score-sim [ER]

we still see some residual moveout in the angle gather. The left panel of Figure 7 shows the input two traces (third and sixteenth).

After running the algorithm we obtained the score and similarity matrices seen in Figure 6. Note how the structure of the similarity matrix to the previous example (Figure 4). The score matrix and the corresponding maximum has the shape that we would anticipate. It is generally diagonal with some deviations. The output two traces appear to be better aligned (the right panel of Figure 7), but the difference isn't as reduced as we would hope. Our belief is this caused by a poor stretching algorithm.

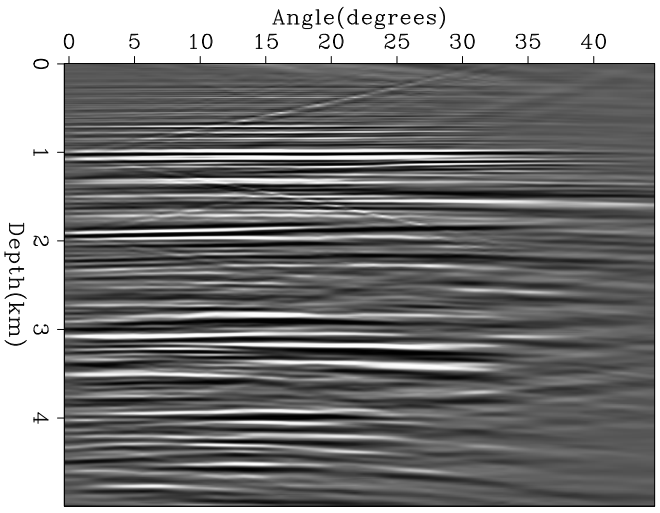


Figure 5: The CRP gather used for trace alignment. The third and 16th trace were used. bob2-big.gather [ER]

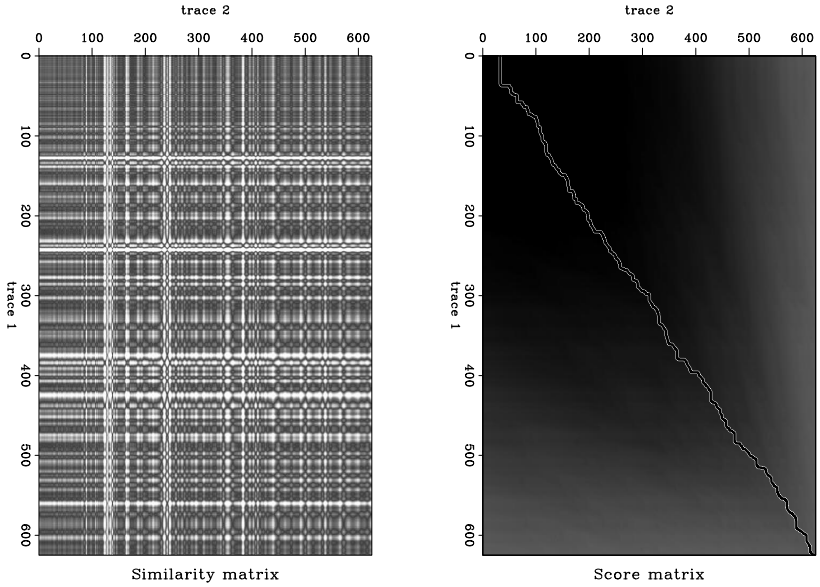


Figure 6: The right plot is the score matrix using two traces from a CRP gather and the left panel is the similarity matrix. Axes labels refer to time sample numbers (not seconds). bob2-big.score-sim [ER]

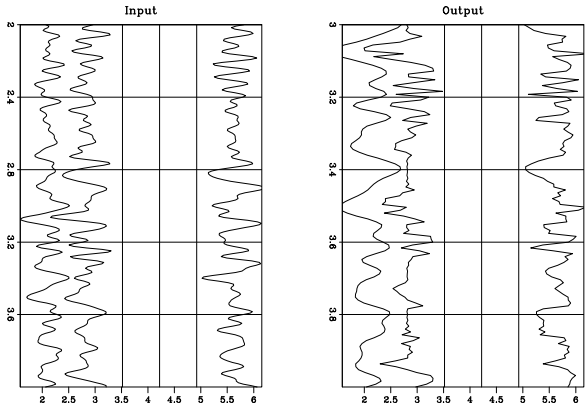


Figure 7: The left plot shows the two input traces and the right plot the traces after alignment in the window from three to four seconds. The third trace in each display shows the difference. bob2-big.in-out [ER]

CONCLUSIONS

We have shown that a robust and efficient algorithm originally developed for protein sequence alignment can be applied to the pairwise alignment of seismic traces. This has been demonstrated, in its simplest form, by application to synthetic and real seismic data.

With further work, this approach may supply a general tool for nonlinear alignment of seismic traces for use in processing and interpretation.

ACKNOWLEDGMENTS

The first author would like to thank J. Claerbout and B. Biondo for the opportunity to visit SEP during May and June of 2002.

REFERENCES

Darby, E. K., and Neidell, N. S., 1966, Application of dynamic programming to the problem of plane-wave propagation in a layered medium: Geophysics, **31**, 869–883.

Gazdag, J., and Sguazzero, P., 1985, Migration of seismic data by phase shift plus interpolation, *in* Gardner, G. H. F., Ed., Migration of seismic data: Society Of Exploration Geophysicists, 323–330.

Karp, R. M. Cs 294-7 lecture notes:. http://www.icsi.berkeley.edu/ epxing/lecture7.pdf, 2000.

Martinson, D. G., and Hopper, J. R., 1992, Nonlinear seismic trace interpolation: Geophysics, **57**, 136–145.

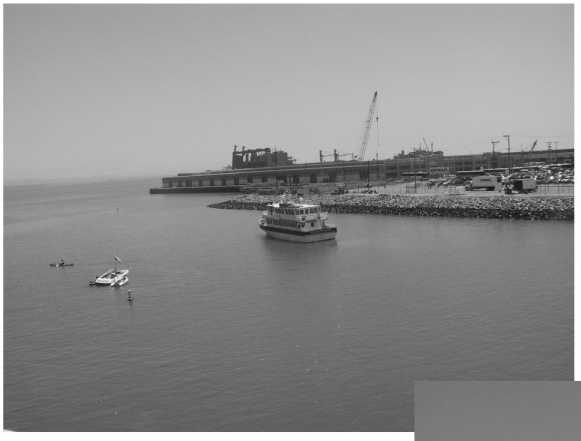
Martinson, D. G., Menke, W., and Hopper, J. R., 1982, An inverse approach to signal correlation: Journal of Geophysics Research, **87**, 4807–4818.

Needleman, S., and Wunsch, C., 1970, A general method applicable to the search for similarities in the amino acid sequence of two proteins: J. Mol. Biol., **48**, 443–453.

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

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

SchneiderW. A., J., Ranzinger, K., and Balch, A., 1992, A dynamic programming approach to first arrival traveltime computation in media with arbitrarily distributed velocities: Geophysics, **57**, 39–50.




PACIFIC BELL PARK


YAHOO!


YAHOO!
MINE

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 181–195

Simultaneous estimation of two slopes from seismic data, applied to signal/noise separation

Morgan Brown[1]

ABSTRACT

I present an efficient new approach to simultaneously estimate two slopes from seismic data. I employ a Newton iteration to overcome the problem's nonlinearity. In spite of my method's theoretical inability to handle aliased data, it robustly estimates two independent slopes in many circumstances. I apply my method to the problem of signal/noise separation on synthetic and real data examples. The estimated slopes provide approximate inverse signal and noise covariance operators good enough to obtain an excellent separation, with only a limited amount of prior information required.

BACKGROUND

For practical purposes, seismic data consists locally of the superposition of n plane waves. Disregarding aliasing effects and the data's wavelet, the slopes of the n plane waves fully and uniquely parameterize the data locally. Claerbout (1992)[2] casts the problem of single-slope estimation as a linear, univariate optimization problem. Fomel (2000; 2001b; 2001a) extends the problem to the estimation of two slopes, and utilizes the estimated slopes for the interpolation of missing data and signal/noise separation. He iteratively solves a linearization of a nonlinear problem, and applies a model regularization term to enforce smoothness of the estimated slopes.

In this paper, I present another method for solving the two-slope estimation problem. It is a nonlinear extension of Claerbout's methodology, and differs from Fomel's in the sense that it is a strictly local method. Because mine is a local method, it runs much faster than Fomel's. Theoretically, my method is sensitive to aliased data, unlike Fomel's. Like Fomel's, the estimated slope depends on the starting guess. The existence of local minima appears to be an inherent weakness of the two-slope estimation problem in general.

Fomel successfully applies "plane-wave destructor" filters, derived from estimated slopes, to the signal/noise separation problem. Analogously, I use the estimated slopes to construct "steering filters" of a form derived by Clapp et al. (1997). Like Fomel, I find that when the signal and noise slopes are too similar, my method converges to (incorrect) local minima, unless the slope estimation is "guided" with a prior model of the signal or the noise. Using

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

[2]section 4.5, *Dip Picking Without Dip Scanning*

this constrained approach, I obtain excellent separation results on three different real data examples. Most encouragingly, in all cases, very simple, easily-obtained prior models sufficed.

THE METHOD

Claerbout (1992) casts the problem of slope estimation as a univariate optimization problem, based on the observation that the partial differential equation

$$\left(\frac{\partial}{\partial x} + p\frac{\partial}{\partial t}\right) u(t,x) \tag{1}$$

is zero-valued if the wavefield $u(t,x)$ consists only of plane waves with time slope, or "stepout," p. Claerbout also notes that a cascade of two PDEs annihilates data consisting of plane waves with two slopes, p_1 and p_2. The analog to equation (1) is:

$$\left(\frac{\partial}{\partial x} + p_1\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial x} + p_2\frac{\partial}{\partial t}\right) u(t,x), \tag{2}$$

or after expansion,

$$\left(\frac{\partial^2}{\partial x^2} + (p_1+p_2)\frac{\partial^2}{\partial x\,\partial t} + p_1 p_2\frac{\partial^2}{\partial t^2}\right) u(t,x). \tag{3}$$

Discretizing the problem

Claerbout approximates the derivatives of equation (1) with 2x2 finite difference stencils. Assuming that the grid spacing in both the t and x directions are unity:

$$\frac{\partial}{\partial x} \approx 0.5 * \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}, \qquad \frac{\partial}{\partial t} \approx 0.5 * \begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix}. \tag{4}$$

By convolving together these first-order stencils, we can construct appropriate finite-difference stencils to approximate the second-order differential operators of equation (3):

$$\frac{\partial}{\partial x} * \frac{\partial}{\partial x} = \frac{\partial^2}{\partial x^2} \approx 0.25 * \begin{bmatrix} -1 & 2 & -1 \\ -2 & 4 & -2 \\ -1 & 2 & -1 \end{bmatrix} \tag{5}$$

$$\frac{\partial}{\partial t} * \frac{\partial}{\partial x} = \frac{\partial^2}{\partial x\,\partial t} \approx 0.25 * \begin{bmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{bmatrix} \tag{6}$$

$$\frac{\partial}{\partial t} * \frac{\partial}{\partial t} = \frac{\partial^2}{\partial t^2} \approx 0.25 * \begin{bmatrix} -1 & -2 & -1 \\ 2 & 4 & 2 \\ -1 & -2 & -1 \end{bmatrix} \tag{7}$$

The stencils of equations (5)-(7) are convolved with the data, $\mathbf{u}$. For simplicity, we can define the following notation:

$$\frac{\partial^2}{\partial x^2} * \mathbf{u} = \mathbf{D}_{xx}; \qquad \frac{\partial^2}{\partial x\,\partial t} * \mathbf{u} = \mathbf{D}_{xt}; \qquad \frac{\partial^2}{\partial t^2} * \mathbf{u} = \mathbf{D}_{tt}, \tag{8}$$

and rewrite equation (3) in matrix form:

$$\mathbf{r} = \begin{bmatrix} \mathbf{D}_{xx} & \mathbf{D}_{xt} & \mathbf{D}_{tt} \end{bmatrix} \begin{bmatrix} 1 \\ p_1 + p_2 \\ p_1 p_2 \end{bmatrix}. \tag{9}$$

The vector $\mathbf{r}$ has the same dimension as the data, $\mathbf{u}$. If the data consists only of plane waves with slopes p_1 and p_2, then equation (9) predicts values of $\mathbf{u}$ from nearby values of $\mathbf{u}$. If the data's slopes change in time and space, however, equation (9) is valid only across local "patches" of the data. We can rewrite equation (9) to reflect this fact:

$$\mathbf{r} = \begin{bmatrix} \mathbf{D}^1_{xx} & \mathbf{D}^1_{xt} & \mathbf{D}^1_{tt} \\ \vdots & \vdots & \vdots \\ \mathbf{D}^n_{xx} & \mathbf{D}^n_{xt} & \mathbf{D}^n_{tt} \end{bmatrix} \begin{bmatrix} 1 \\ p_1 + p_2 \\ p_1 p_2 \end{bmatrix} \tag{10}$$

Equation (10) denotes the convolution of the respective finite-difference stencils over a data patch of size n, where n may be as large as the entire data, or as small as 3×3

While it is tempting to make a change of variables ($a = p_1 + p_2, b = p_1 p_2$) and treat equation (10) as a linear relationship, I have found that this approach produces trivial coupled estimates of the true slopes. This problem is inherently nonlinear.

Dip Estimation

To estimate two local slopes p_1 and p_2, we treat vector $\mathbf{r}$ in equation (10) as a familiar prediction error, and find the p_1 and p_2 which minimize the squared norm of the prediction error. First we define the following shorthand:

$$\sum \mathbf{D}_{xx}\mathbf{D}_{xt} = \sum_{i=1}^{n} \mathbf{D}^i_{xx}\mathbf{D}^i_{xt}.$$

Expanding $\mathbf{r}^T\mathbf{r}$ from equation (10) and collecting terms yields a nonlinear function of p_1 and p_2, which we denote $Q(p_1, p_2)$:

$$\begin{aligned} Q(p_1,p_2) &= \sum \mathbf{D}^2_{xx} + p_1 \cdot 2 \sum \mathbf{D}_{xx}\mathbf{D}_{xt} + p_2 \cdot 2 \sum \mathbf{D}_{xx}\mathbf{D}_{xt} \\ &+ p_1 p_2 \cdot \left(2 \sum \mathbf{D}_{xx}\mathbf{D}_{tt} + 2 \sum \mathbf{D}^2_{xt} \right) + p_1^2 \cdot \sum \mathbf{D}^2_{xt} + p_1^2 p_2 \cdot 2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} \\ &+ p_2^2 \cdot \sum \mathbf{D}^2_{xt} + p_1 p_2^2 \cdot 2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1^2 p_2^2 \cdot \sum \mathbf{D}^2_{tt}. \end{aligned} \tag{11}$$

To find the least-squares-optimal p_1 and p_2, we compute the partial derivatives of $Q(p_1, p2)$, set them equal to zero, and solve a system of two equations.

$$\begin{aligned} \frac{\partial Q(p_1,p_2)}{\partial p_1} = f(p_1,p_2) &= \sum \mathbf{D}_{xx}\mathbf{D}_{xt} + p_2 \sum \mathbf{D}_{xx}\mathbf{D}_{tt} + p_2 \sum \mathbf{D}^2_{xt} \\ &+ 2 p_1 p_2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} + p_2^2 \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1 p_2^2 \sum \mathbf{D}^2_{tt} = 0 \end{aligned} \tag{12}$$

$$\begin{aligned} \frac{\partial Q(p_1,p_2)}{\partial p_2} = g(p_1,p_2) &= \sum \mathbf{D}_{xx}\mathbf{D}_{xt} + p_1 \sum \mathbf{D}_{xx}\mathbf{D}_{tt} + p_1 \sum \mathbf{D}^2_{xt} \\ &+ 2 p_1 p_2 \sum \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1^2 \mathbf{D}_{xt}\mathbf{D}_{tt} + p_1^2 p_2 \sum \mathbf{D}^2_{tt} = 0 \end{aligned} \tag{13}$$

I use Newton's method for two variables to compute the optimal slopes by updating estimates of p_1 and p_2 with the following iteration:

$$p_{1,k+1} = p_{1,k} + \frac{-f(p_{1,k},p_{2,k})g_{p_2}(p_{1,k},p_{2,k}) + f_{p_2}(p_{1,k},p_{2,k})g(p_{1,k},p_{2,k})}{f_{p_1}(p_{1,k},p_{2,k})g_{p_2}(p_{1,k},p_{2,k}) - f_{p_2}(p_{1,k},p_{2,k})g_{p_1}(p_{1,k},p_{2,k})} \quad (14)$$

$$p_{2,k+1} = p_{2,k} + \frac{-f_{p_1}(p_{1,k},p_{2,k})g(p_{1,k},p_{2,k}) + f(p_{1,k},p_{2,k})g_{p_1}(p_{1,k},p_{2,k})}{f_{p_1}(p_{1,k},p_{2,k})g_{p_2}(p_{1,k},p_{2,k}) - f_{p_2}(p_{1,k},p_{2,k})g_{p_1}(p_{1,k},p_{2,k})} \quad (15)$$

The estimated slopes at iteration k are $p_{1,k}$ and $p_{2,k}$. $f_{p_1}(p_{1,k},p_{2,k})$ is, for example, the partial derivative of $f(p_1,p_2)$ with respect to p_1. While intimidating, equations (14) and (15) result simply from the inversion of a 2-by-2 matrix of second derivatives (of $Q(p_1,p_2)$), the so-called Hessian matrix. Since, the partial derivatives of f and g are non-constant, the problem is non-quadratic, which implies that Newton's method may diverge for certain initial guesses $(p_{1,0},p_{2,0})$, and furthermore, may converge to a local minimum. In practice, however, the method converges to machine precision within 3-5 iterations.

SLOPE ESTIMATION TESTS

Figures 1 and 2 illustrate tests of the my nonlinear two-slope estimation algorithm. The "textures" (Brown, 1999) were computed by constructing nonstationary steering filters (the 9-point Lagrange filter derived by Clapp et al. (1997)) with the estimated slopes, and then using those filters to deconvolve random noise. The textures, used also by Fomel (2000; 2001b), provide a quick check of the accuracy of the estimated slopes.

In Figure 1, a simple crossing-plane-wave dataset is tested. The slope panels shown have been smoothed with a sliding weighted mean filter (4-by-4 analysis window). The program used to compute the slopes also computes the weights, which are either 1 or 0. If the estimated slopes at a single point in (t,x) are equal, then the result is assumed to be trivial and the weight at that point is set to 0. Otherwise, the weight is set to 1.

The textures in Figure 1 illustrate that the estimated slopes are not totally accurate. The steep positive slope in particular seems smaller than the true positive slopes in the data, while the shallower negative slope seems better represented.

Figure 2 illustrates a more difficult test dataset, a "CMP gather" overlain by upward-sloping linear "noise." Notice that some regions of the data contain either one signal or the other. My slope estimation program, through the use of mask operators, allows the user to specify regions where only one slope is present in the data. In those regions, I use Claerbout's (1992) univariate "puck" method to estimate that single slope.

We notice some discontinuity in the textures in Figure 2 at one-slope/two-slope boundaries in the data. My two-slope algorithm slightly underestimates the positive "noise" slope, while in some sections of the data, it overestimates the magnitude of the "CMP gather" slope. Still, the general trend of both slopes honors those present in the data.

At the right and bottom edges of the estimated slopes in Figure 2, notice the constant-valued regions. Because the finite-difference templates of equation (7) run off the right and

bottom edges of the data, the slope cannot (easily) be computed in these regions. In this case, the slope remains unchanged from the starting guesses, which in this case were -0.5 and 0.5. We expect the estimated slopes to exhibit some sensitivity to starting guess. I have experimented qualitatitvely, and indeed found some sensitivity, though it is not generally severe.

SIGNAL/NOISE SEPARATION

Consider the recorded data to be the simple superposition of signal and noise events: $\mathbf{d} = \mathbf{s} + \mathbf{n}$. The so-called *Wiener estimator* is a filter, which when applied to the data, produces an optimal (least-squares sense) estimate of the embedded signal (Castleman, 1996). For the special case of uncorrelated signal and noise, the frequency response of this filter is

$$\mathbf{H} = \frac{\mathbf{P_s}}{\mathbf{P_n} + \mathbf{P_s}}, \tag{16}$$

where $\mathbf{P_s}$ and $\mathbf{P_n}$ are the signal and noise power spectra, respectively. Define operators $\mathbf{N}$ and $\mathbf{S}$, as convolution with filters which decorrelate the unknown noise $\mathbf{n}$ and signal $\mathbf{s}$, respectively. Brown and Clapp (2000), for example, show that the following least-squares optimization problem is approximately equivalent to Wiener estimation:

$$\begin{aligned} \mathbf{Nn} &\approx 0 \\ \epsilon\mathbf{Ss} &\approx 0 \\ \text{subject to} &\leftrightarrow \mathbf{d} = \mathbf{s} + \mathbf{n} \end{aligned} \tag{17}$$

Equation (17) is a regularized linear least-squares problem. The scalar parameter ϵ is related to the data's signal-to-noise ratio.

The conceptual model of seismic data as n locally-crossing plane waves lends itself well to parameterization by a few parameters. The multidimensional prediction-error filter (PEF) is a particularly popular option (see, for example, (Claerbout, 1998)). Estimated by autoregression against the data, the PEF encodes hidden multiplicity in the data with a few filter coefficients. It has the approximate inverse spectrum of the data from which it was estimated.

By using a model of the noise to obtain a nonstationary noise PEF and deconvolving a PEF estimated from the data by the noise PEF to obtain a signal PEF (Spitz, 1999), many authors have solved equation (17) to successfully separate coherent noise from signal (Spitz, 1999; Brown et al., 1999; Clapp and Brown, 2000; Brown and Clapp, 2000; Guitton et al., 2001).

As noted by Fomel (2000), however, the considerable amount of parameter tuning required to create stable nonstationary PEFs (a requirement for the deconvolution step) remains a significant obstacle to their use in industrial-scale processing environments.

If the signal and noise consist of distinct slopes everywhere, then it is in theory possible to implicitly separate signal from noise in the slope domain with a two-slope estimation algorithm. Fomel uses estimated slope to construct plane-wave destructor filters which are used

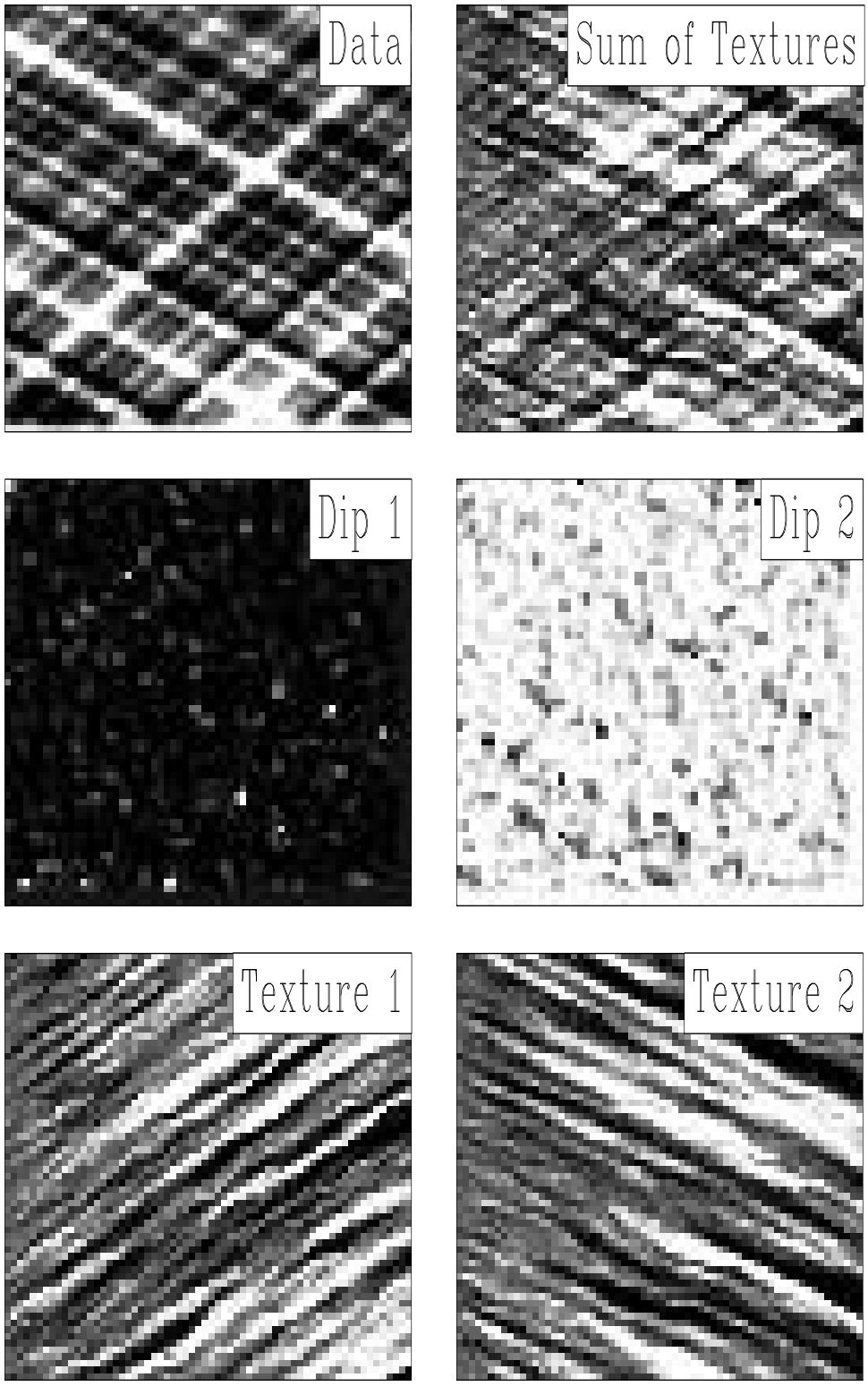


Figure 1: Test of my nonlinear two-slope estimation algorithm on a simple synthetic test case. morgan1-pucknl.dumb2 [ER,M]

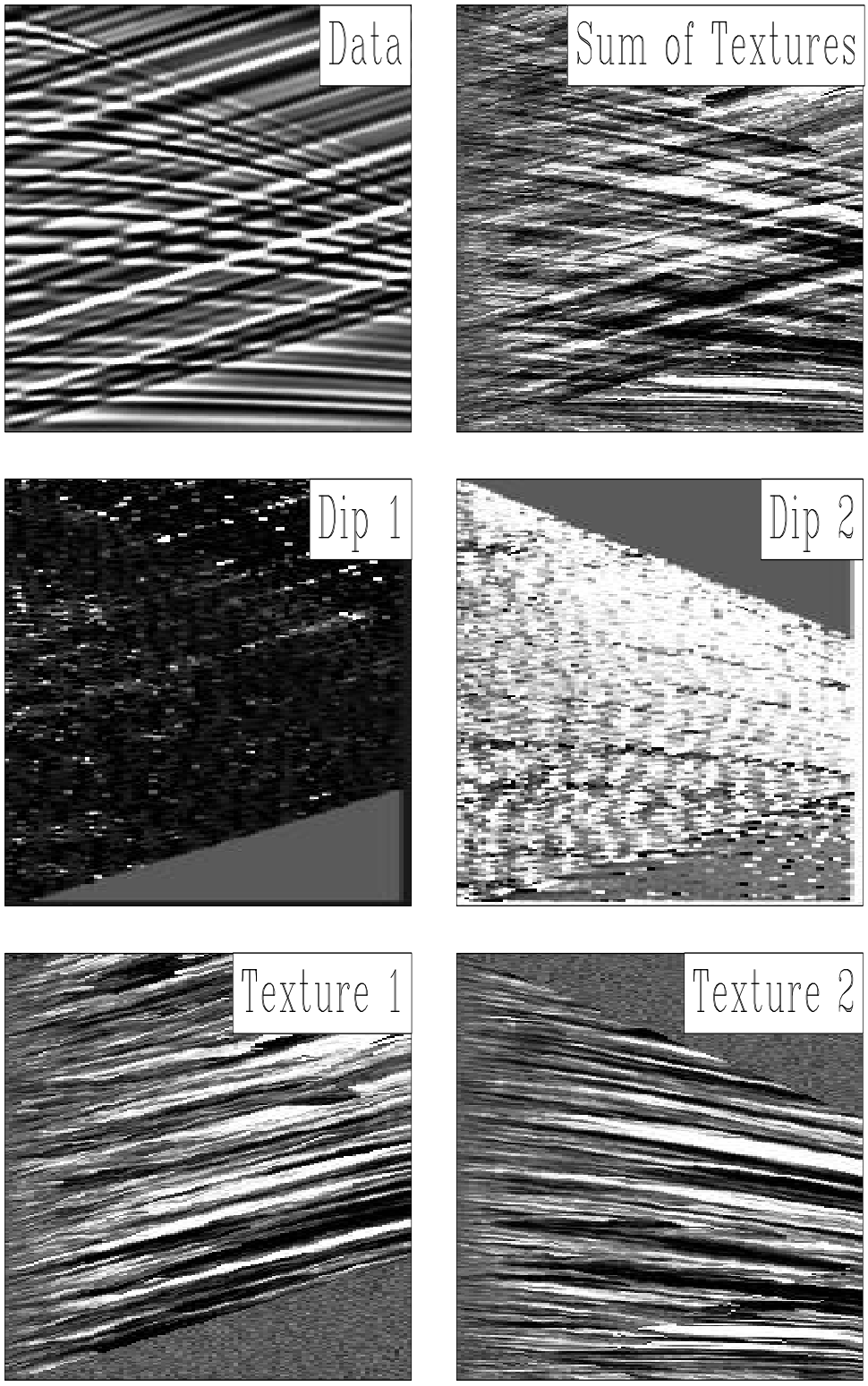


Figure 2: Test of nonlinear 2-slope "puck" algorithm on a more realistic synthetic test. morgan1-pucknl.dumb [ER,M]

directly as **N** and **S** in equation (17), without any deconvolution. The filters are guaranteed stable and insensitive to spatially aliased data. Fomel obtains an independent estimate the noise slope from a prior noise model, and then fixes the noise slope as the signal slope is estimated.

I take a slightly different tack at the problem. Like Fomel, I use my two-slope estimation technique to directly obtain signal and noise slope estimates. I also exploit a prior noise model and also a prior signal model, in cases where the signal is simpler to model than the noise. Most importantly, I find that very simple, easily-obtainable signal or noise models suffice. To overcome aliasing, I apply normal moveout (NMO) to the data. Rather than plane-wave destructor filters, I (again) use 9-point Lagrange steering filters derived by Clapp et al. (1997).

Constrained signal/noise separation results

I tested my signal/noise separation approach on three real data examples: two common-midpoint (CMP) gathers infested with multiples and a 2-D slice from a 3-D terrestrial shot gather with strong ground roll. In each case, I generate prior models of the signal and the noise (details described below), which are used in the following two-stage procedure:

1. Use single-slope estimation to obtain noise/signal slope from the prior noise/signal model.
2. Use two-slope estimation, with the slope from step 1 fixed, to estimate the other slope.
3. Initialize signal and noise steering filters and solve equation (17).

The first test, shown in Figure 3, was performed on a CMP gather from a 2-D dataset acquired by WesternGeco in the Gulf of Mexico. Characterized by strong water-bottom and top-of-salt multiple reflections, this dataset was the focus of the 1997 SEG multiple attenuation workshop. To model the signal and noise, I adopted a simple approach: given a random zero-offset section and a stacking velocity, apply inverse NMO and inverse NMO for first-order water-bottom multiples (Brown, 2002) to create a signal and noise models, respectively.

The panels on the top row of Figure 3 have been NMO-corrected with the stacking velocity, to facilitate comparison. Estimated primary reflections should be flat. We can see from the estimated signal and noise panels that my approach has produced an excellent separation result. Many totally obscured primaries now appear from beneath the multiple train. To dealias the data, I applied NMO with water velocity (4900 ft/sec). The effects of this step can be seen in the panels on the prior models and estimated dips panels of Figure 3. In this case, I fixed the *signal* slope, not the noise slope.

The method performed less impressively at near offsets, where the signal and noise dips both tend to 0. In these regions, the inversion simply "splits the difference," according to the ϵ parameter in equation (17). I conducted a second multiple separation test on CMP gather taken from the "Mobil AVO" dataset (Lumley et al., 1994). The gather is characterized by a strong train of first order water-bottom multiples, as well as strong primary events partially

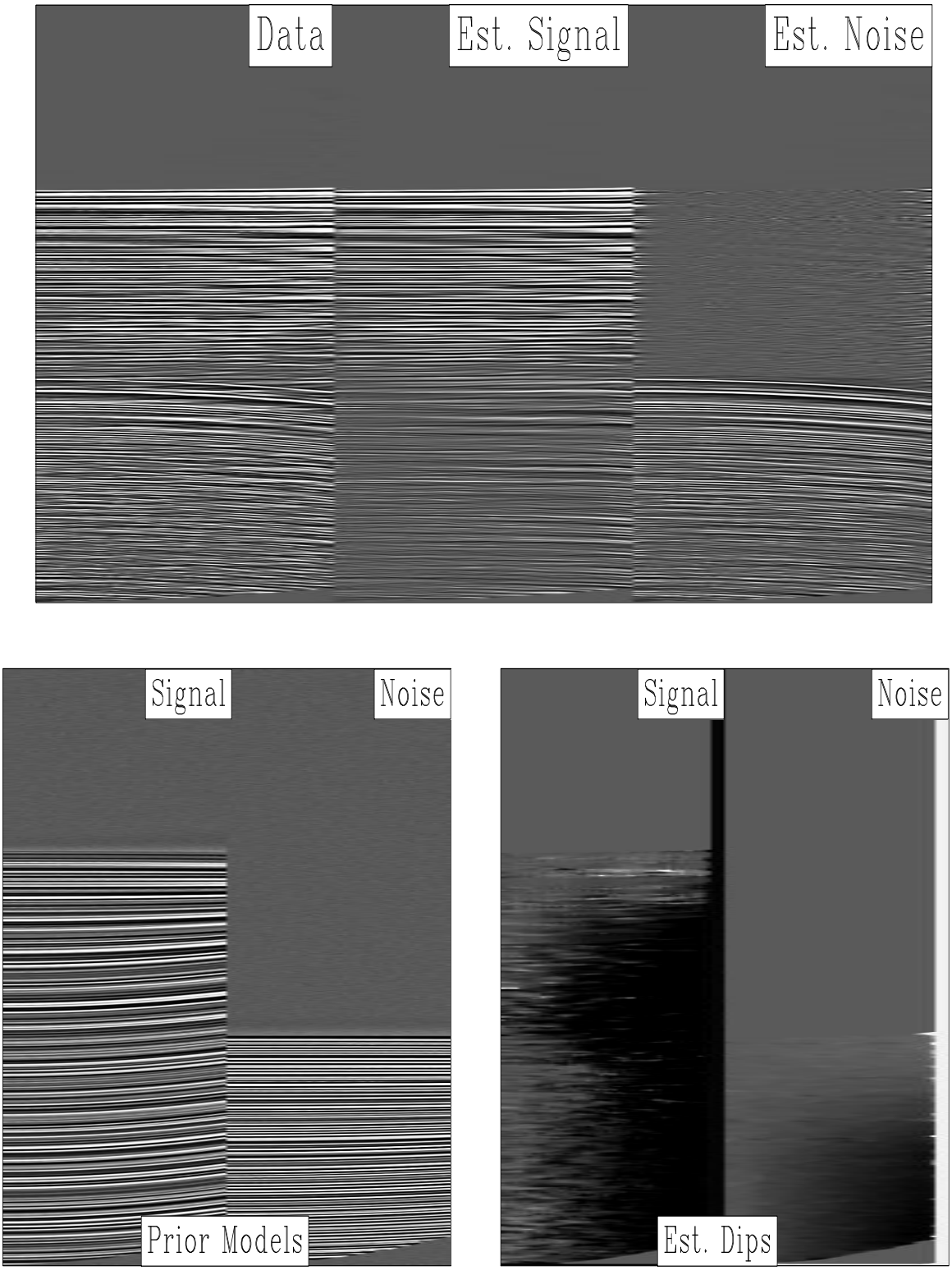


Figure 3: Signal/Noise separation tested on Gulf of Mexico CMP gather. Top row, left to right: original data, estimated signal, estimated noise. Bottom row, left to right: Prior signal and noise models, estimated signal and noise dips. morgan1-sn.gulfcmp [ER,M]

concealed under the noise. These deep primaries provide an excellent benchmark to test the signal preservation characteristics of my method.

The separation results are shown in Figure 4. As with the previous example (Figure 3), inverse NMO was used to create the prior models, and the prior signal slope was fixed in the slope estimation step. Again, the separation results are quite impressive. The embedded primary events appear to be perfectly preserved, and most of the reverberations are segregated to the noise panel. We expect the same near-offset behavior as the previous example, though the results in this region look plausible. Like before, the data were dealiased with an NMO-correction of 2.0 km/sec. My final test was conducted on a 2-D receiver line extracted from a 3-D shot gather acquired by Saudi Aramco. While the ground roll looks impossibly strong to conceal any extractable information, there is indeed no shortage of primaries under the noise cone.

The signal and noise are approximately, but not perfectly, separable in temporal frequency (Brown et al., 1999). To obtain an approximate noise model, I applied a lowpass filter with a cutoff of 10 Hz. Conversely, I obtained a signal model by applying a highpass filter with a 35 Hz cutoff. Unlike the previous two examples, I treated the noise slope as fixed in the slope estimation step. In this case, the noise is simpler than the signal. We have confidence in the noise slope; everything else is treated as signal.

While the results are not as impressive, they are good nontheless. It is not difficult to find many coherent primary events that have been unmasked from under the strong ground roll. Notice that near zero offset, some noise has leaked into the signal model. Although the data were again dealiased with an NMO correction (decreasing velocity), the noise is still spatially aliased at far offsets, a fact confirmed by a look at the estimated noise slope. Th separation results are visibly compromised in those regions. Furthermore, a persistent "ringing" is present around zero offset. Fomel (2000, 2001b,a) solved the problem by supplementing the noise decorrelation filter [**N** in equation (17)] with a 3-point notch filter.

CONCLUSIONS

I presented a nonlinear two-slope extension to Claerbout's (1992) single-slope estimation method. Compared to Fomel's approach, (2000; 2001b; 2001a), my method is less accurate and theoretically sensitive to aliasing, but also faster. I demonstrated very encouraging results in the application of signal/noise separation, where the two slopes in the data were associated to signal and noise steering filters.

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.

Brown, M., Clapp, R. G., and Marfurt, K., 1999, Predictive signal/noise separation of groundroll-contaminated data: SEP–**102**, 111–128.

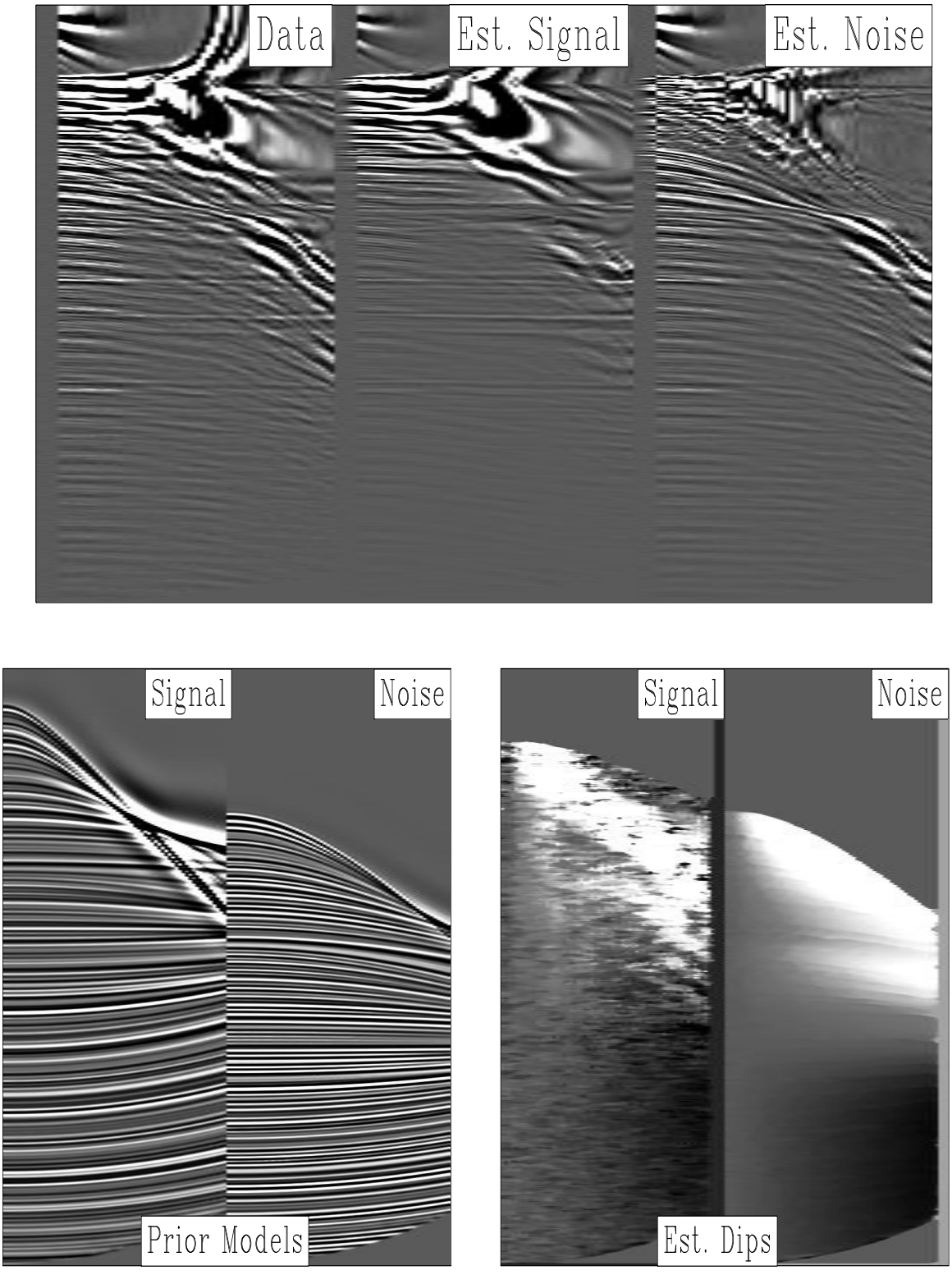


Figure 4: Signal/Noise separation tested on Mobil AVO CMP gather. Top row, left to right: original data, estimated signal, estimated noise. Bottom row, left to right: Prior signal and noise models, estimated signal and noise dips. morgan1-sn.haskreal [ER,M]

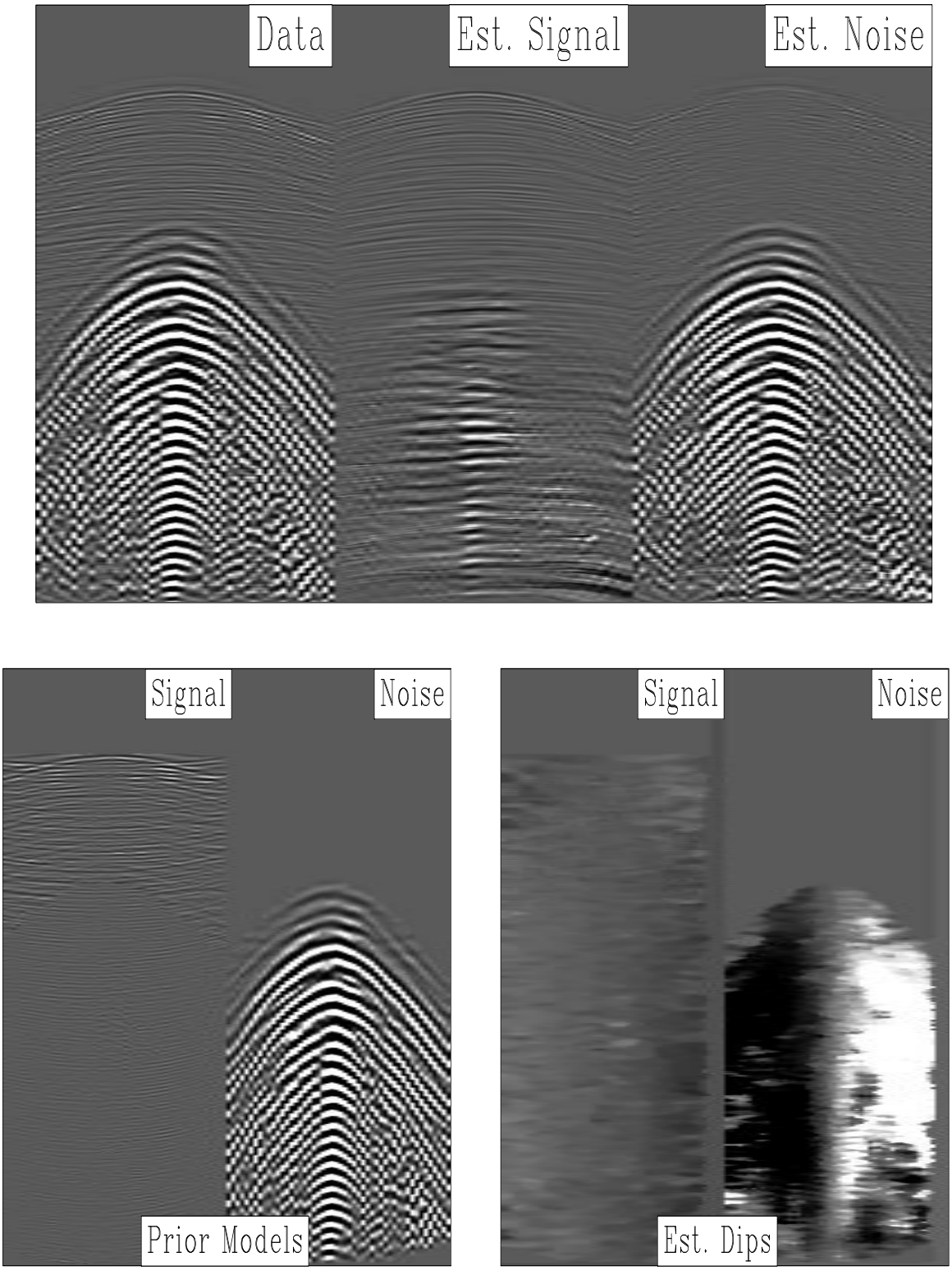


Figure 5: Signal/Noise separation tested on Saudi ground roll data. Top row, left to right: original data, estimated signal, estimated noise. Bottom row, left to right: Prior signal and noise models, estimated signal and noise dips. morgan1-sn.dune [ER,M]

Brown, M., 1999, Texture synthesis and prediction error filtering: SEP–**100**, 211–222.

Brown, M., 2002, Least-squares joint imaging of primaries and multiples: SEP–**111**, 33–47.

Castleman, K. R., 1996, Digital image processing: Prentice-Hall.

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

Claerbout, J. Geophysical estimation by example: Environmental soundings image enhancement:. http://sepwww.stanford.edu/sep/prof/, 1998.

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

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

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

Fomel, S., 2001a, Seismic data interpolation and noise attenuation with plane-wave destructor filters: 71st Ann. Internat. Mtg, Soc. Expl. Geophys., Expanded Abstracts, 1985–1988.

Fomel, S., 2001b, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.

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.

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

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–112, September 27, 2002, pages 195–213

Can we make genetic algorithms work in high-dimensionality problems?

Gabriel Alvarez[1]

ABSTRACT

In this paper I compare the performance of a standard genetic algorithm versus a micro-genetic algorithm for matching a randomly-generated seismic trace to a reference trace with the same frequency spectrum. A micro genetic algorithm evolves a very small population that must be restarted whenever the population loses its genetic diversity. I show that the micro-genetic algorithm is more efficient in solving this problem in terms of improved rate of converge, especially in the first few generations. This characteristic may make the method useful for locating the most promising valleys in the search space which can then be searched with more traditional gradient-based methods. An additional benefit is a significant reduction in the number of evolution parameters that needs to be adjusted making the method more easy to use.

INTRODUCTION

The kind of optimization problems that we usually face in exploration geophysics are non-linear, high-dimensional, with a complex search space that may be riddled with many local minima or maxima. Usually our first line of attack is linearization of the problem around some given smooth, "easily" computed initial model. Seismic tomography is a good example of the success of this approach. There are many cases, however, where linearization is impractical or undesirable and the full non-linear problem must be solved. Broadly speaking, there are two ways to attack these kind of problems: deterministic search methods, for example non-linear conjugate gradient, quasi-Newton methods or Levenberg-Marquardt method (Gill and Murray, 1981), and stochastic search methods such as Montecarlo, simulated annealing and genetic algorithms (Davis, 1987; Goldberg, 1989a). Deterministic methods are attractive because they are natural extensions of familiar linear methods and because, in certain applications, they can be made to run extremely fast. The downside is the need to compute first and/or second order derivatives of the cost function and the dependence of the solution (and sometimes even the convergence of the method itself) on a suitable starting point. In other words, deterministic methods are extremely efficient at locating the bottom of the valley, provided they start the search somewhere inside the valley. This is a serious shortcoming since in many problems locating the valley that contains the minimum may be a problem as difficult as locating the

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

minimum itself. We could say that deterministic methods are poor at "exploration" (locating the best valleys) but are very good at "exploitation" (given the valley, locating its floor).

Stochastic methods, on the other hand, perform a much more exhaustive search of the model space but are not as good at exploiting the early results of the search. It appears that a hybrid method combining the strengths of both techniques would be the best choice. The situation is not so clear cut, however, because it may be difficult to identify the most promising valleys and it may also be difficult to compute the derivatives of the cost function. In geophysics such a hybrid approach between a genetic algorithm and conjugate gradient was used to solve the problem of estimating velocities from refraction seismic data, although the results were not conclusively better than those obtained by the genetic algorithm alone (Boschetti, 1995).

An interesting alternative is the use of the so-called micro-genetic algorithms (Krishnakumar, 1989) which aim at improving the relatively poor exploitation characteristic of the genetic algorithms without affecting their strong exploration capabilities. In this paper I compare the results of applying both a standard and a micro-genetic algorithm to the problem of matching a seismic trace. This is part of the more interesting problem of inverting a zero-offset trace for interval velocities addressed in a companion paper in this report (Alvarez, 2002). Figure 1 shows the input sub-sampled sonic log and the corresponding reference seismic trace obtained by a simple computation of the normal incidence reflection coefficients assuming no multiples, no absorption and no noise.

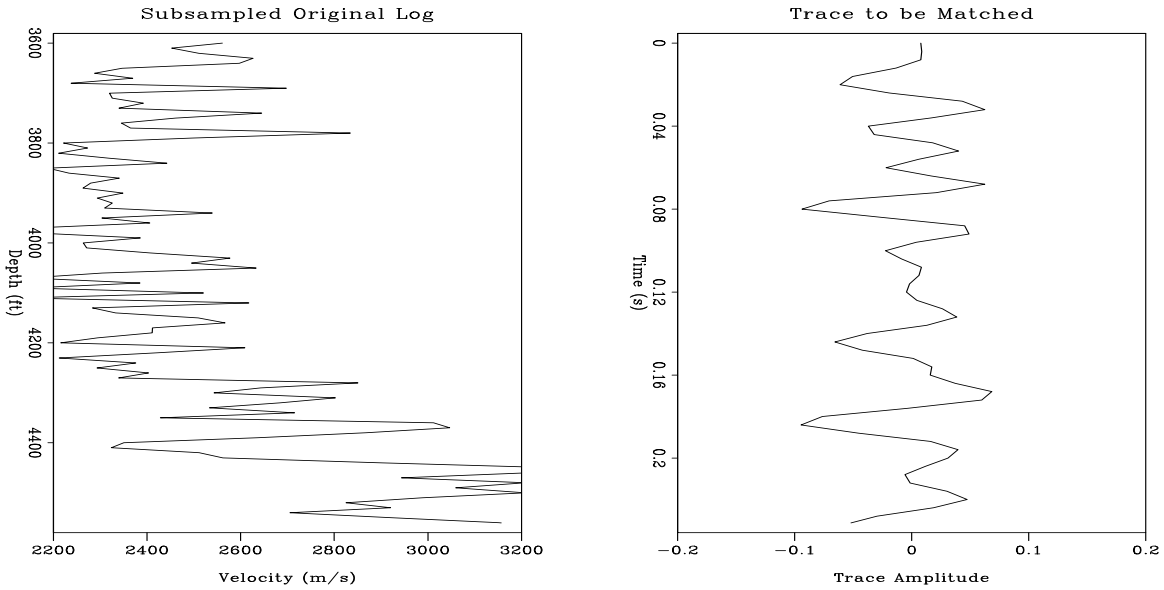


Figure 1: Left, sub-sampled sonic log used to generate the synthetic seismic trace on the right gabriel1-input [ER]

STANDARD GENETIC ALGORITHM

A genetic algorithm is an optimization method inspired by evolution and survival of the fittest. A trial solution to the problem is constructed in the form of a suitably encoded string of model parameters, called an *individual*. A collection of individuals is in turn called a *population*. There are several considerations and choices to be made in order to implement a suitable solution to an optimization problem using genetic algorithms. A full description of all the practical details is outside the scope of this paper, and some of them are a matter of active research (Gen and Cheng, 2000; Haupt and Haupt, 1998; Falkenauer, 1998; Beasley et al., 1993). In Appendix A I give a brief description of the most relevant issues of genetic algorithm optimization as used in this study. In particular, I describe model-parameter encoding as well as standard and non-standard operators (selection, jump and creep mutation, crossover, elitism and niching), fitness function and convergence criteria.

Parameter Selection

The first step, of course, is to select the *model* parameters related to the "physics" of the problem, their ranges (maximum and minimum values) and their required resolution. Then it is time to choose the *evolution* parameters actually related to the genetic algorithm itself. The performance of a genetic algorithm to solve a particular optimization problem depends critically on the choice of its evolution parameters that must be fine-tuned to that problem as much as possible. In general it is difficult to give hard and fast rules that may work with a wide range of applications, although some guidelines exist (Goldberg and Richardson, 1987; Goldberg, 1989b; Goldberg and Deb, 1991). For this problem, I choose the evolution parameters one by one starting with the most critical and working my way down to the least critical. Once a particular parameter is selected it is kept constant in the tests to select the remaining parameters. I do so because it would be nearly impossible to test all possible combinations. The reader may find useful to refer to Appendix A for a description of the evolution parameters themselves.

Model Parameter Encoding

For the present application I use binary encoding of the 99 *model* parameters (not to be confused with the *evolution* parameters that control the inner workings of the genetic algorithm) representing potential solutions to the problem (Alvarez, 2002). I represent a model parameter with 10 bits so that there are $2^{10} = 1024$ possible values for each model parameter and a total of 1024^{99} possibilities for the entire search space. A completely exhaustive search of the model space would therefore be very difficult if not impossible.

Population Size

Ideally, the population size should be large enough to guarantee adequate genetic diversity yet small enough for efficient processing. In particular, the number of cost-function evaluations is

proportional to the population size. Equation 1 corresponds to Goldberg's criterion (Goldberg, 1989b) of increasing the population size exponentially with the increase in the number of model parameters (assuming binary encoding).

$$npopsize = order[(l/k)(2 * *k)] \tag{1}$$

In this equation, l is the number of bits in the chromosome and k the order of the schemata of interest (schemata is plural for schema [2]). This criterion, however, may result in populations too large when the number of model parameters and so the length of the chromosome is large. In this case, for example, with 99 model parameters, each encoded with 10 bits, even for a relatively low-order schema of 5 bits the population size would have to be larger than 3000 individuals. Such large populations may require many cost-function evaluations and a lot of memory. Most applications reported in the literature use population sizes between 50 and 200 individuals and I am unaware of any reported use of a population larger than 1000 individuals. For the purpose of the present application I decided to try a relatively wide number of population sizes using for the other *evolution* parameters "reasonably" standard choices. For example, I used uniform crossover with a probability of 0.6, jump mutation probability equal to $1/npopsize$ (large populations have larger genetic diversity and so less need for jump mutation) and creep mutation probability equal to the number of bits per model parameter times the jump mutation rate. Also, for this first set of tests I allowed both niching (sharing) and elitism of the best individual.

The top panels of Figure 2, from left to right, show a comparison of convergence rate as a function of number of generations for population sizes of 50, 100, 200 individuals whereas the bottom panels show similar curves for population sizes of 250, 500 and 1000 individuals. As expected, the larger populations produce better convergence after a fixed number of generations. This is not a fair test, however, since the number of cost-function evaluations is proportional to the population size. Figure 3 shows the same curves for a fixed number of function evaluations. Clearly, there is a practical range of optimum population sizes about 200 or 250 individuals. For all the following tests I used a population of 200 individuals. Figure 4 shows a comparison of the reference trace (solid line) with the inverted traces (dotted line) obtained with each population in the same order as that in Figure 3. The difference in the match of the traces is not so impressive because in all cases a good solution is eventually found for all populations. The smallest population, however, does show a poorer match than the others.

Selection Mechanism

I used a tournament selection in which each parent is the best fit of two individuals picked at random from the population. This technique has the advantage of applying significant selection pressure while avoiding the pitfalls of fitness ordering or ranking (Falkenauer, 1998).

[2] A schema is a similarity template describing a subset of strings with similarities at certain string positions. For example, the schema 0*1 matches the two strings 001 and 011.

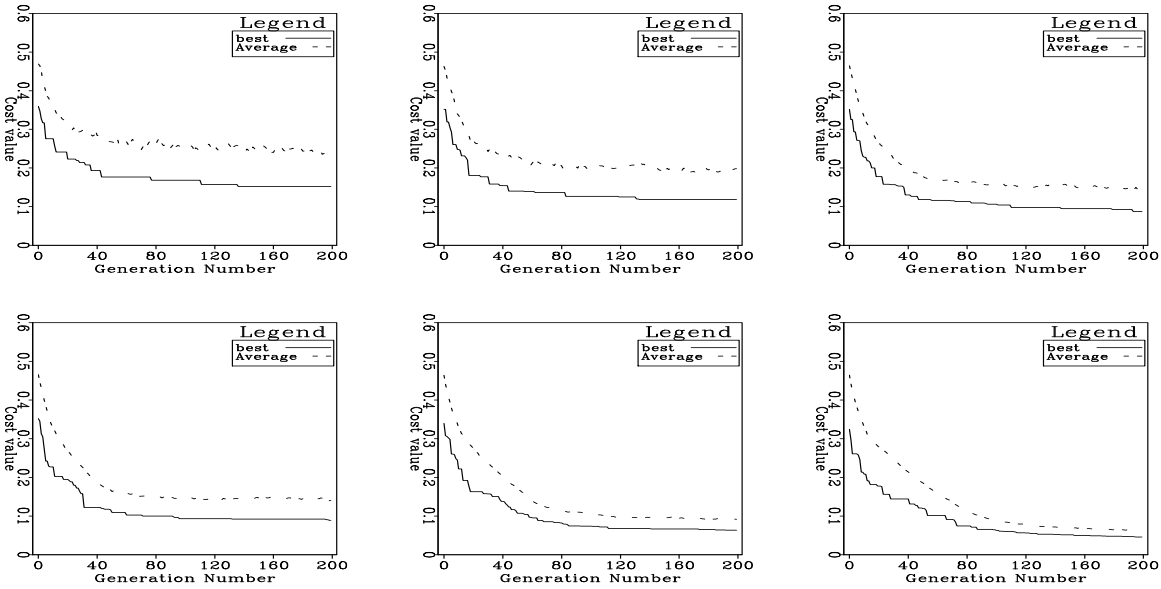


Figure 2: Comparison of convergence rates for different population sizes keeping the number of generations constant. Top row populations 50, 100 and 200 individuals. Bottom row, populations of 250, 500 and 1000 individuals. gabriel1-SG_compare_pop_sizes1 [ER]

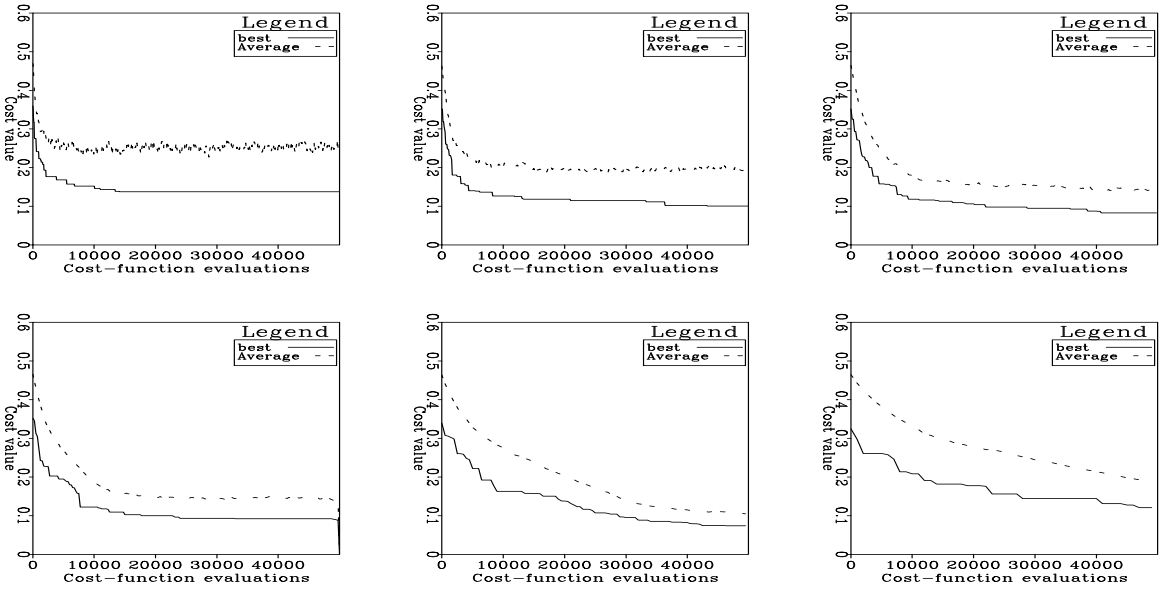


Figure 3: Same as in Figure 2 but in terms of number of function evaluations rather than number of generations. gabriel1-SG_compare_pop_sizes2 [ER]

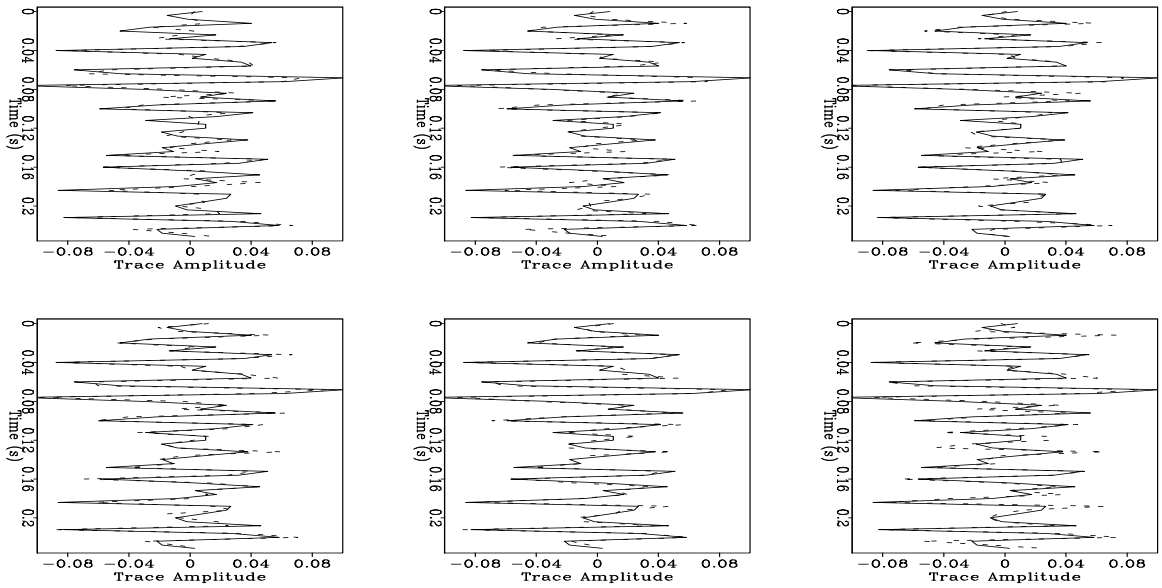


Figure 4: Comparison of trace match for different population sizes keeping the number of function evaluations constant. Continuous line is the reference trace and dotted line the inverted trace. Top row, populations of 50, 100 and 200. Bottom row, populations of 250, 500 and 1000. gabriel1-SG_compare_tr_pop_sizes2 [ER]

Crossover Rate

Crossover is by far the most important evolution operation. I tested single-point and uniform crossover with a crossover probability of 0.6. The population size was chosen to be 200 and the algorithm was run for 250 generations. The jump mutation was set at 0.005 and the creep mutation at 0.05. Elitism of the best individual as well as niching was allowed. The top panels in Figure 5 show the convergence rate of the two cases, whereas the bottom panels show the corresponding traces. In this case uniform crossover (right panel) performs a little better since it reaches a lower cost-value after the allowed number of cost-function evaluations. Using uniform crossover, I tried six different values of crossover probability: 0.5, 0.6, 0.7, 0.8, 0.9 and 1.0. I tried this large range of values because crossover rate is a particularly important evolution parameter. The top panels of Figure 6 show the comparison of the convergence rates for crossover rates of 0.5, 0.6 and 0.7 whereas the bottom panels show the same curves for crossover rates of 0.8, 0.9 and 1.0. The results are surprisingly similar, although it appears that the smaller crossover rates produce faster initial convergence, and so I chose a crossover rate of 0.6 for the remaining tests.

Mutation

As mentioned in Appendix A, there are two types of mutation operators: the standard jump mutation that acts on the chromosome (binary representation of the individual, sometimes called *genotype*) and creep mutations that act on the decoded individual, sometimes called

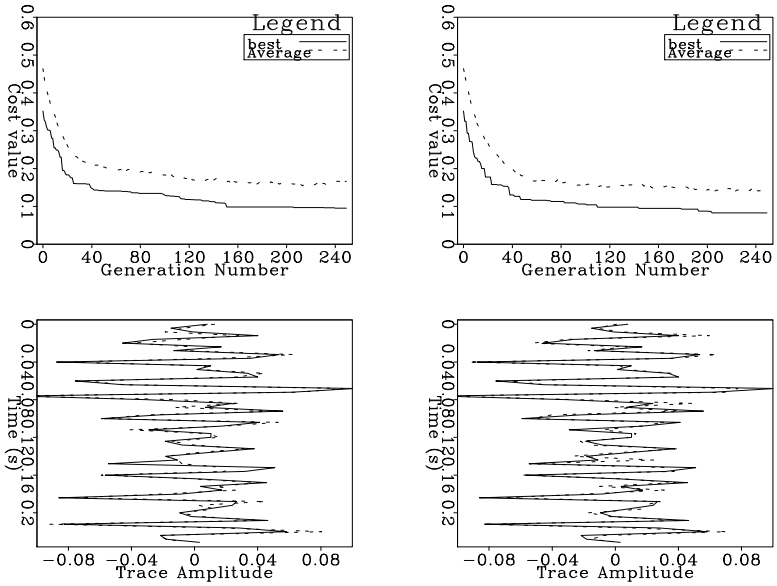


Figure 5: Comparison of convergence rates for two types of crossover: single-point (left) and uniform (right). Top panels are convergence rates whereas bottom panels are trace match with continuous line representing the reference trace and dotted line the inverted trace. gabriel1-SG_compare_crossover1 [ER]

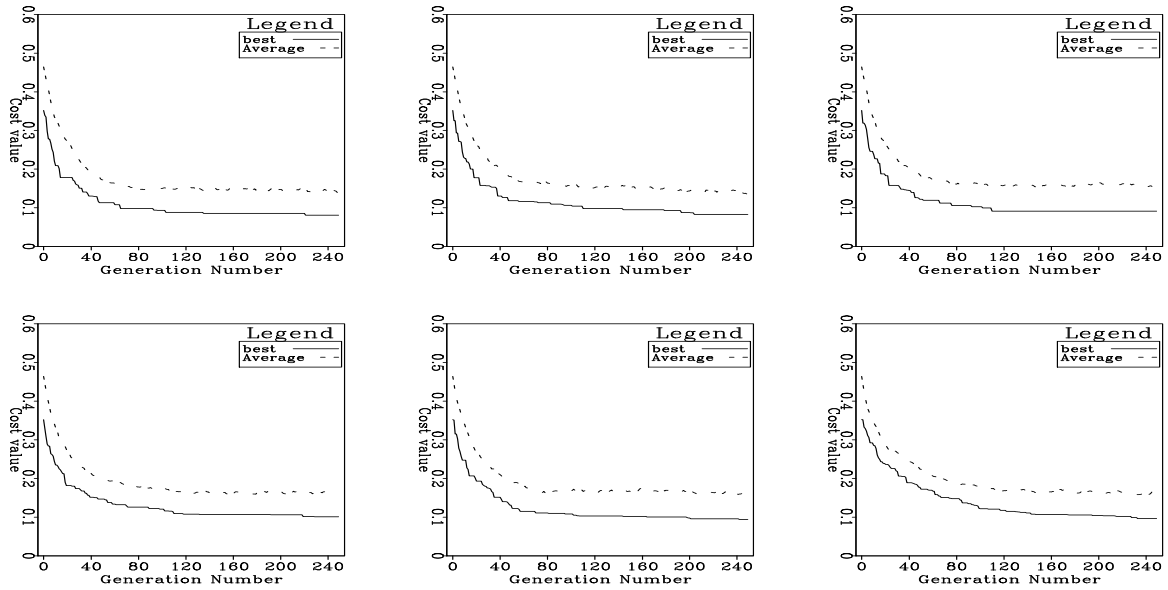


Figure 6: Comparison of convergence rates for different crossover rates. Top panels correspond to crossover rates of 0.5, 0.6 and 0.7 whereas bottom panels correspond to crossover rates of 0.8, 0.9 and 1.0. gabriel1-SG_compare_crossover2 [ER]

phenotype. In any case, the mutation probabilities are expected to be low, since high values may cause strong disruption of promising schemata and therefore steer the algorithm away from the most promising regions of the search landscape. The top panels of Figure 7 show a comparison of convergence rates for three values of jump mutation probability 0.002, 0.004 and 0.008, without creep mutations. In this case a jump mutation of 0.002 is the best although a value of 0.005 would have been expected from the rule of thumb of the inverse of the population size. The bottom panels of Figure 7 show the effect of creep mutations with probabilities of 0.02, 0.04 and 0.08. Again, it seems that a small mutation is actually best. For the remaining I used creep mutation with probability 0.02.

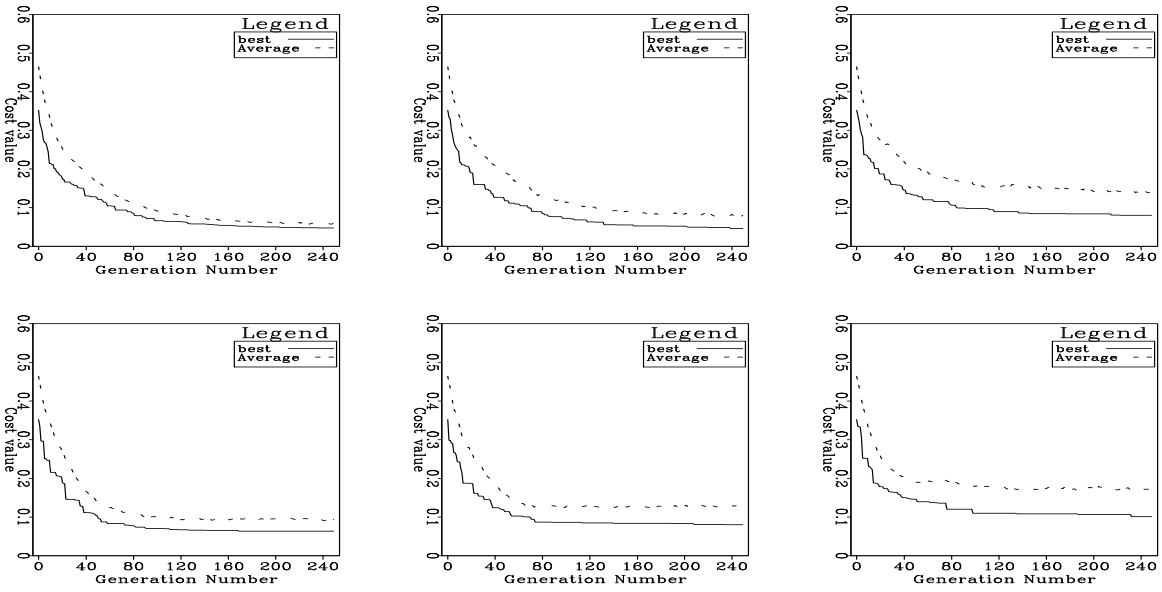


Figure 7: Comparison of convergence rates for different jump and creep mutation rates. Top panels for jump mutation rates of 0.002, 0.004 and 0.008. Bottom panels for creep mutation rates of 0.02, 0.04 and 0.08. gabriel1-SG_compare_mutation1 [ER]

Other options

Finally, I tested the inclusion of niching (Goldberg and Richardson, 1987) and elitism. The top panels of Figure 8 compare the convergence rates with niching and elitism (left)i, only elitism (middle), and only niching (right). The bottom panels show a similar comparison for the trace match. All the other evolution parameters were chosen according to the previous analysis. It seems that the lack of niching makes the convergence a little slower in the first iterations (compare the left and middle panels) but makes it faster after about the 40th generation. The lack of elitism, on the other hand, makes convergence a little erratic (compare left and right panels) since we are not guaranteed to go to a better-fit best individual in any generation compared with the previous one. Elitism was thus included and niching excluded in the evolution parameters of the final optimum standard genetic algorithm.

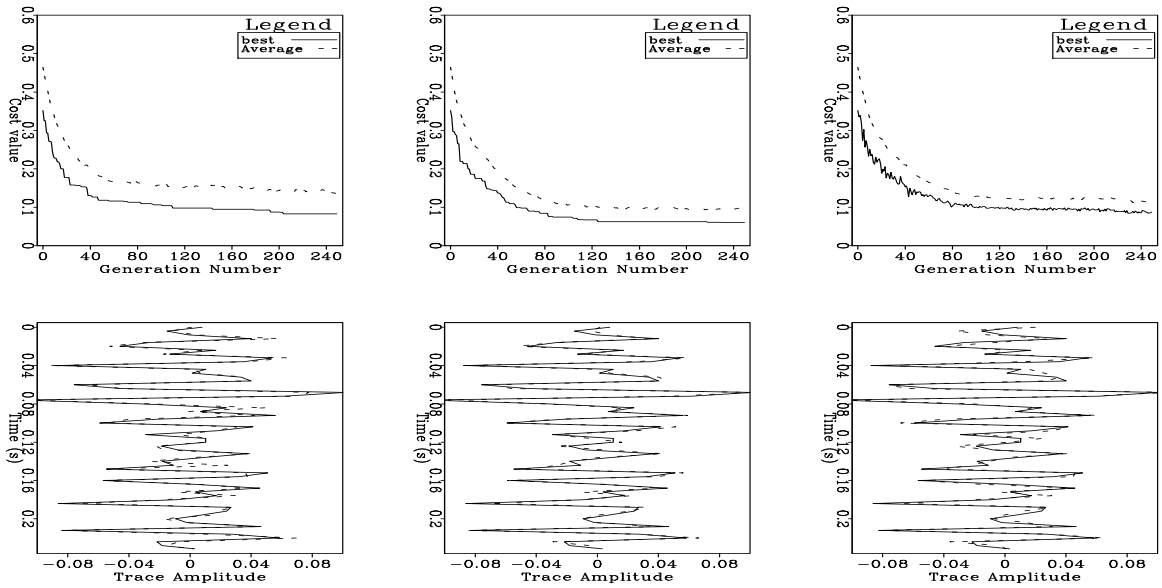


Figure 8: Comparison of results for the inclusion of niching and elitism. Top panel convergence rates: left with niching and elitism, middle with elitism only and right with niching only. Bottom panels show similar comparison for the trace match with continuous line representing the reference trace and dotted line the inverted trace. gabriel1-SG_compare_other1 [ER]

Parameter Summary

For easy reference, Table 1 shows the evolution parameters selected as optimum from the previous tests. The resulting genetic algorithm will be compared with a micro-genetic algorithm described in the next section.

Table 1: Summary of the optimum evolution parameters for the standard genetic algorithm

Population size	200	Crossover rate	0.6
Mutation rate	0.002	Creep mutation rate	0.02
Elitism	Yes	Niching	No
Selection strategy	Tournament	Number of children	1

MICRO-GENETIC ALGORITHM

When dealing with high dimensionality problems, it may be difficult or too time consuming for all the model parameters to converge within a given margin of error. In particular, as the number of model parameters increases, so does the required population size. Recall that large population sizes imply large numbers of cost-function evaluations. An alternative is the use of

micro-genetic algorithms (Krishnakumar, 1989), which evolve very small populations that are very efficient in locating promising areas of the search space. Obviously, the small populations are unable to maintain diversity for many generations, but the population can be restarted whenever diversity is lost, keeping only the very best fit individuals (usually we keep just the best one, that is, elitism of one individual). Restarting the population several times during the run of the genetic algorithm has the added benefit of preventing premature convergence due to the presence of a particularly fit individual, which poses the risk of preventing further exploration of the search space and so may make the program converge to a local minimum. Also, since we are not evolving large populations, convergence can be achieved more quickly and less memory is required to store the population.

Selection of Evolution Parameters for Micro-GA

In principle, micro-genetic algorithms are similar to the standard genetic algorithm described in the previous section, in the sense of sharing the same evolution parameters and similar considerations. There is, however, an important distinction: since new genetic material is introduced into the population every time the algorithm is restarted, there is really no need for either jump or creep mutation. Also, elitism is required, at least every time the population is restarted, otherwise the algorithm would lose its exploitation capability. I have also found that the algorithm is much less sensitive to the choice of evolution parameters compared with the standard genetic algorithm. In particular, population sizes of 5 to 7 with crossover rates of 0.8 to 0.95 give very good results. The top panels of Figure 9 shows a comparison of convergence rates for populations of 3, 5 and 7 individuals. It seems clear that 5 individuals is the best. This result agrees with Carroll's who employed micro-GAs to optimize an engineering problem (Carroll, 1996). The bottom panels show a comparison of convergence rates for populations of 5 individuals and crossover rates of 0.7, 0.9 and 1.0. It seems that 0.9 is the best crossover rate, although further tests showed that 0.95 gave even better results and therefore that value was chosen for the remaining tests. Figure 10 shows the results in terms of trace match. Again, it is apparent that a population of 5 and a crossover rate of 0.9 are optimum. In particular, note how a uniform crossover of 1.0 (bottom right panel) is far too disruptive.

Summary of Evolution Parameters for Micro-GA

Table 2 shows a summary of the evolution parameters selected for the micro-GA for the current application.

COMPARISON OF STANDARD AND MICRO-GA

Having chosen the evolution parameters that provided the best results for both the standard and the micro-GA (Tables 1 and 2), I will now compare the performance of the two in solving the current problem.

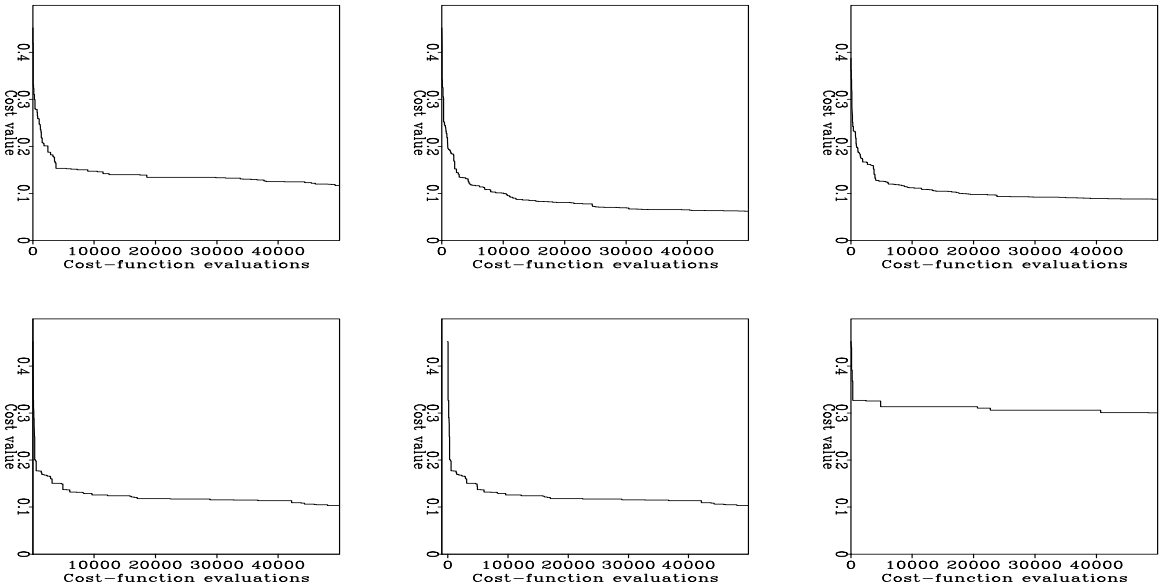


Figure 9: Comparison of convergence rates for different options of the micro genetic algorithm. Top panel population rates of 3, 5 and 7 (from left to right). Bottom panels, with population size of 5 and crossover rate of 0.7, 0.9 and 1.0. gabriel1-MG_compare1 [ER]

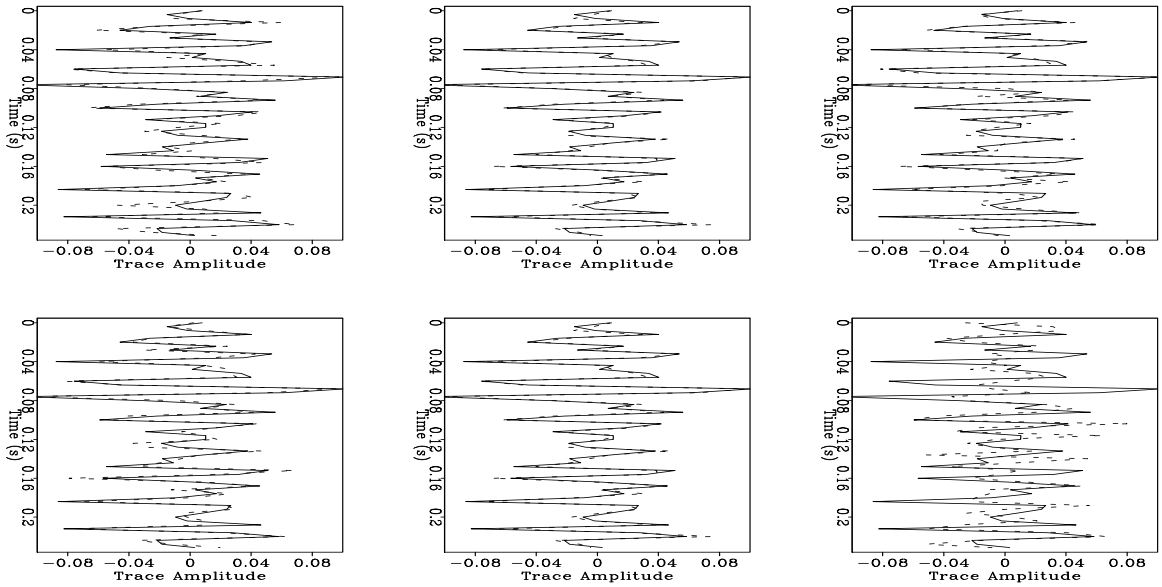


Figure 10: Comparison of trace match for different options of the micro genetic algorithm. Continuous line is the reference trace and dotted line is the inverted trace. Top panel population rates of 3, 5 and 7 (from left to right). Bottom panels, with population size of 5 and crossover rate of 0.7, 0.9 and 1.0. gabriel1-MG_compare2 [ER]

Table 2: Summary of the optimum evolution parameters for the micro-genetic algorithm

Population size	5	Crossover rate	0.95
Jump mutation rate	0.0	Creep mutation rate	0.0
Elitism	Yes (best individual only)	Niching	No
Selection strategy	Tournament	Number of children	1

The top panels of Figure 11 show a comparison of convergence rates between the standard genetic algorithm (left) and the micro-genetic algorithm (right). The difference in convergence rate is in the first generations is impressive. For example, the micro-genetic algorithm would have essentially converged after 2000 cost-function evaluations, whereas the standard genetic algorithm would take almost 10000 cost-function evaluations to reach the same convergence level. If enough iterations (generations) are allowed both algorithms will converge to essentially the same result. The bottom panels in Figure 11 show the corresponding trace match. The differences are not too great because both algorithms were essentially run to convergence.

CONCLUSIONS AND FUTURE WORK

The results of this test are encouraging because micro genetic algorithms show a much faster rate of convergence than standard genetic algorithms in the solution of this simple, relatively high-dimensional problem. At the very least micro genetic algorithms could be run for a few generations and use the results as starting points for gradient-based methods.

From the point of view of the genetic algorithm inversion, some lessons have been learned after extensive testing of the evolution parameters. Firstly, using a micro-genetic algorithm with uniform cross-over without mutation emerges as the best option for this problem (as opposed to a standard genetic algorithm with single-point cross-over and jump and creep mutation). Secondly, a micro-genetic-algorithm population of 5 individuals with a cross-over probability of 0.95 seems to be optimum for this problem.

An important issue to be further analyzed is that of the multi-modality of the search space. In this case it is clear that there is a single global minimum, namely recovering the original trace sample-by-sample. However, I have found that once I get close enough to this global minimum it takes a large number of iterations to escape local minima (many traces “almost fit” exactly the original). My present convergence criteria do not allow for checking of convergence of individual model parameters so I have to investigate alternative options.

Another important issue has to do with the convenience of working with the model parameters directly in their floating-point representation rather than the standard binary encoding used here. This approach has the advantage of not requiring a resolution limit on the model parameters.

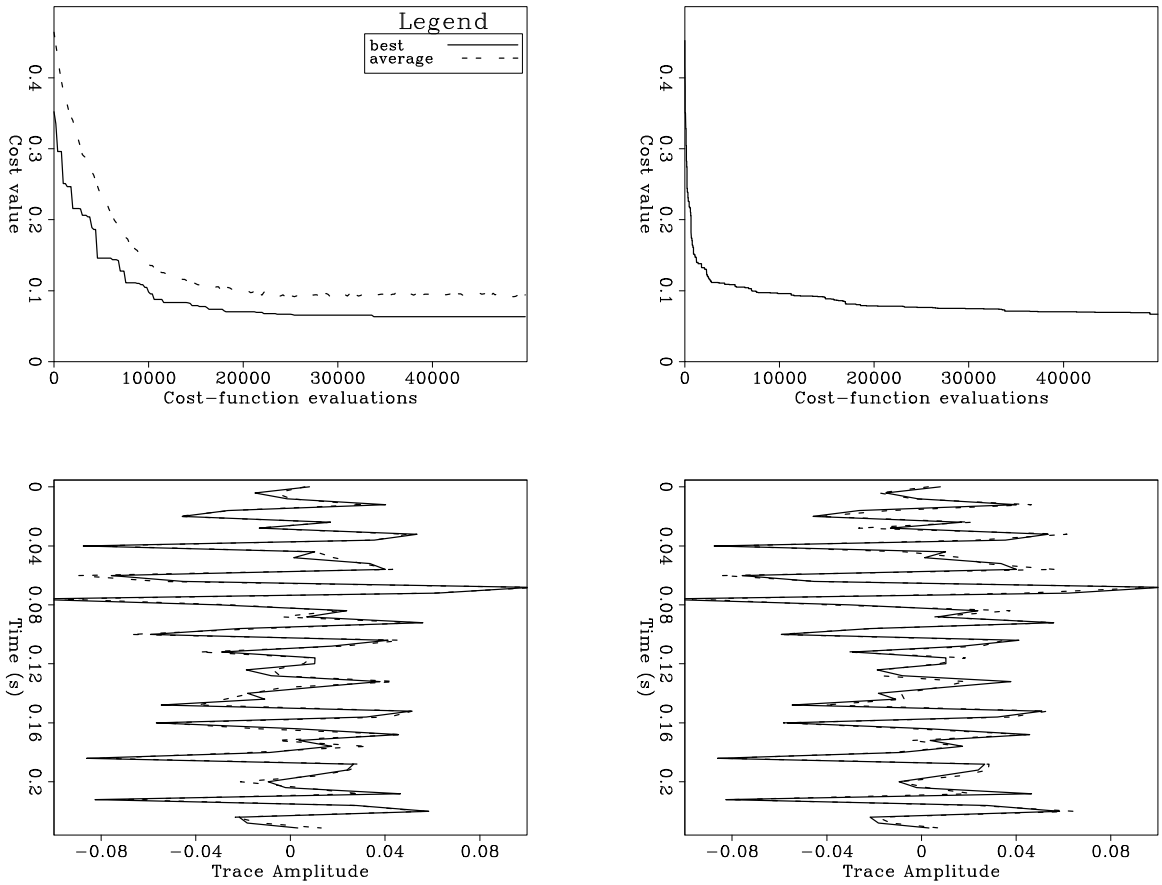


Figure 11: Comparison between the standard and the micro genetic algorithm. Top panels convergence rates for standard genetic algorithm (left) and micro genetic algorithm (right). Bottom panels show the corresponding trace match with continuous line representing the original trace and the dotted line the inverted ones. gabriel1-final_comparison [ER]

APPENDIX A: REVIEW OF GENETIC ALGORITHMS

In this appendix I briefly review some of the terms and issues related to genetic algorithms in optimization. More detailed accounts can be found in (Goldberg, 1989a; Haupt and Haupt, 1998; Falkenauer, 1998; Gen and Cheng, 2000).

Model Parameter Encoding

The choice of representation of the problem in terms of how many model parameters, their encoding, range of values and required resolution is perhaps the most important decision we face when using genetic algorithms. In particular, we must decide whether to use "direct" representation of the problem (in terms of floating point numbers, for example) or to "encode" the solution in terms of a suitable "alphabet," usually binary. In his pioneering work in genetic algorithms Holland employed the binary representation to prove his schemata theorem which

provides the theoretical foundation for the workings of the genetic algorithm (Holland, 1975; Goldberg, 1989a). This theorem proves that short, low order schemata are more likely to be preserved by the evolution process in contrast to long high order schemata which are more likely to be disrupted by crossover and mutation.[3] Short, low order schemata, therefore, are likely to end up associated with highly fit individuals, that is, with the best solutions to the problem. Therefore, if we can encode our model parameters in such a way that the promising short low-order schemata are produced, we may achieve a faster convergence and obtain a better solution. If we use binary encoding, in order to have short, low-order schemata, the model parameters must be suitably encoded with related model parameters being put close together in the binary string representing an individual (Goldberg, 1989a). The problem is that in general we may not know before hand which parameters are related to others or to what extent they are related. Therefore, in general it is difficult to establish the order of the predominant schemata in a given encoding of the model parameters.

Basic Operators

A genetic algorithm optimization begins by generating a population of randomly computed individuals within the constraints imposed on the model parameters. Three basic operators: reproduction (selection), crossover and mutation are used to "evolve" the solution from one generation (iteration) to the next.

Selection

Each trial solution (individual) is assigned a figure of merit that represents how good a solution it is according to the fitness function (cost function or objective function) of the problem. The most fit individuals (lowest cost-values for minimization problems) are given a higher probability of mating in order to produce the next generation. There are several ways to select the "parents" for mating, such as random pairing, roulette wheel, rank weighting and tournament selection (Haupt and Haupt, 1998). Whatever the selection method, the net effect is to skew the next generation towards the most fit individuals, that is, towards the most promising regions of the search space. It is still possible and desirable to allow less fit individuals to mate, albeit with a lower probability. This increases the exploration of the search space and helps prevent premature convergence, i.e., convergence to a local minimum.

Crossover

Crossover is the operator actually responsible for the exchange of genetic material between the parents in order to produce their offspring. In the usual case of binary encoding, the simplest crossover operator randomly selects a bit position in the binary string representing an

[3]The order of a schema is the number of fixed positions. For example, the order of 011*1** is 4. The length of a schema is the distance between the first and the last specific string positions. For example, the schema 011**1* has a length of 6-1=5

individual (a chromosome), and then two children are produced by taking the bits to the left of the crossover point in parent one and those to the right in parent two and vice-versa. This is called single-point crossover. More than one point may be selected, two, for example, such that one child consists of the bits from the two points to the ends in one parent and those in between the two points from the other parent. The other child will be similarly produced by exchanging the role of the two parents. In the limit a child may be produced by randomly selecting, for each bit, the value from one or the other parent. This is called uniform crossover. It is important to notice that crossover is not necessarily applied to all couples, but only according to a given probability, usually between 0.5 and 0.9. Also, when using other encodings, for example floating-point numbers, the operator must be adjusted (Haupt and Haupt, 1998).

Mutation

The operator responsible for introducing new genetic material into the population is called mutation (or more precisely jump mutation). With binary encoding, this operator works by randomly selecting a bit and then flipping it. In general, mutation is applied with a relatively low probability, usually less than 0.1. The idea is that although mutations are critical to prevent the population from loosing their genetic diversity (that is, mutations force the algorithm to look into other areas of the search space) they are potentially disruptive causing the algorithm to lose information from a promising search area.

Other operators

There are many other operators that may be used to improve the chances of a fast and accurate convergence of genetic algorithms. The ones used in this study are

- Elitism: This is the operator that promotes, without change, the best individual or individuals of the population to the next generation. These individuals are still eligible for recombination with others to produce the offspring.
- Niching: Also called sharing is the process of sharing genetic material between closely related individuals (Goldberg and Richardson, 1987)
- Creep mutation: Similar to standard mutation but only changes that result in small perturbations of the decoded model parameters (as opposed to their binary representation) are allowed. This can be considered a fine-tuning operator.

Fitness function

The fitness function is the equivalent to the cost function in standard optimization theory. In a minimization problem the fitness of each individual is evaluated and a figure of merit assigned such that the individuals with the smallest cost-function values are considered the most fit.

Convergence

In general, establishing the convergence of a genetic-algorithm optimization may not be an easy matter because we do not know for sure if the algorithm has converged to a local or global minimum. In the first case we would like the algorithm to continue exploring the search space and in the second case we would like the algorithm to stop. There are several different ways in which we can proceed, for example:

- Set a threshold for the minimum cost function that constitutes an acceptable solution and stop the algorithm when this value is reached. The problem with this approach is that in some cases it may be too difficult to establish a priori what a good threshold should be.
- Set a threshold for the *difference* between the best and the average fit individual in the population. This has the advantage of not requiring a priori knowledge of the actual cost function values.
- Set a threshold for the difference between the best individuals of the present and the previous generation or generations. This has the advantage of preventing the algorithm from attempting to refine the search too much by slowly crawling to the bottom of the valley.
- Set a maximum number of generations to be evolved. This is a fail safe criterion which can be useful when comparing the performance of genetic algorithms with different combination of evolution parameters.

In most cases a combination of two or more of these and similar criteria are employed.

REFERENCES

Alvarez, G., 2002, Velocity inversion of a seismic trace with a micro genetic algorithm: SEP–**112**, 213–222.

Beasley, D., Bull, D., and Martin, R., 1993, An overview of genetic algorithms: Part 2, research topics: University Computing, **15**, no. 4, 170–181.

Boschetti, F., 1995, Application of genetic algorithms to the inversion of geophysical data: Ph.D. thesis, The University of Western Australia.

Carroll, D., 1996, Genetic algorithms and optimizing chemical oxygen-iodine lasers: Developments in theoretical and applied mechanics, **18**, 411–424.

Davis, L., 1987, Genetic algorithms and simulated annealing.: Pitman.

Falkenauer, E., 1998, Genetic algorithms and grouping problems: John Wiley and sons.

Gen, M., and Cheng, R., 2000, Genetic algorithms and engineering optimization: John Wiley and sons.

Gill, P., and Murray, W., 1981, Practical optimization: Academic Press.

Goldberg, D., and Deb, K., 1991, A comparative analysis of selection schemes used in genetic algorithms: A comparative analysis of selection schemes used in genetic algorithms:, Morgan Kaufmann, Foundations of Genetic algorithms.

Goldberg, D., and Richardson, J., 1987, Genetic algorithms with sharing for multimodal function optimization: Genetic Algorithms and their application. Proceedings of the Second International Conference on Genetic Algorithms, Morgan Kaufmann Publishers.

Goldberg, D., 1989a, Genetic algorithms in search for optimization and machine learning: Addison-Wesley Pub. Co.

Goldberg, D., 1989b, Sizing populations for serial and parallel genetic algorithms: Proceedings of the Third International Conference on Genetic Algorithms, Morgan Kaufmann Publishers.

Haupt, R., and Haupt, S., 1998, Practical genetic algorithms: John Wiley and sons.

Holland, J., 1975, Adaptation in natural and artificial systems: University of Michigan Press.

Krishnakumar, K., 1989, Micro-genetic algorithms for stationary and non-stationary function optimization.: SPIE: Intelligent control and adaptive systems, **1196**, 289–296.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 213–223

Velocity inversion of a seismic trace with a micro-genetic algorithm

Gabriel Alvarez[1]

ABSTRACT

An ever present goal of seismic processing and inversion is to extract meaningful geologic information from seismic data. The simplest possibility is to invert for seismic impedance, or, considering density constant, invert for velocity. Here I use a micro-genetic algorithm program to achieve this goal. A micro-genetic algorithm is different from a standard genetic algorithm in that it evolves a very small population that must be restarted whenever genetic diversity is lost. I use a real sonic log to compute a synthetic seismic trace and then use that trace as input to the micro-genetic algorithm program to invert for the velocity log. Without further input, the velocity inversion cannot hope to recover the general velocity-depth trend because this information corresponds to the very low frequencies which are absent in the seismic data. In order to achieve a good match between the real and the inverted sonic log, the inversion must be suppled with an estimate of the velocity-depth trend. Here I use a 33-point Savitzky-Golay filter with a sixth-degree polynomial to smooth the velocity log and show that with this trend the velocity log is reasonably well recovered. I used both the L_1 and the L_2 norms of the sample-to-sample difference between the reference and the inverted trace and show that the results with both norms are similar.

INTRODUCTION

Genetic algorithms have always been recognized as powerful tools for inverting complex objective functions with complex constraints either continuous or discrete (Goldberg, 1989; Haupt and Haupt, 1998; Falkenauer, 1998; Beasley et al., 1993a,b) but there is always a question mark about their robustness in handling high-dimensional problems with reasonable accuracy and speed. In a companion paper in this report (Alvarez, 2002) I show that genetic algorithms, and in particular micro-genetic algorithms (Krishnakumar, 1989), can be used to solve a relatively high-dimensional problem. In this paper I give the details of that problem: inverting a seismic data trace for the underlying interval velocities assuming a layer-cake velocity model. This is a relatively simple non-linear problem that has been investigated for at least 20 years (Lindseth, 1982). Several techniques have been proposed and various commercial software packages exist with high levels of sophistication, using deterministic inversion or simulated annealing.

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

If we can use genetic algorithms to efficiently solve this kind of problems, we can then take advantage of the simplicity with which genetic algorithms handle all kinds of continuous or discrete constraints to go beyond the inversion of seismic data for interval velocities. Multiples, wave mode conversion and other such complications may be equally handled.

In this paper I employ a micro-genetic algorithm to invert a synthetic seismic trace for the underlying velocities that produced it. The seismic trace is created from a real well log assuming normal-incidence, a layer-cake velocity model and no multiples or absorption. I show that when the inversion is supplied with an estimate of the velocity-depth trend obtained from the well log with a 33-point, sixth order Savitzky-Golay filter (Press et al., 1992), the inversion yields a very reasonable estimation of the input sonic log.

PREPROCESSING OF THE WELL LOG

The well log used in this project is from Colombia. A 1000-ft long segment of the sonic log from depths 3600 to 4600 feet was selected for the study. This depth interval corresponds to a predominantly shaly sequence with a marked transition to sandstone at the bottom. Originally the sonic log was sampled at 0.5 feet but I sub-sampled it down to 10 feet intervals using Backus average method (Mavko et al., 2000). The sub-sampled log thus contains 99 samples. The left panel of Figure 1 shows the sub-sampled sonic log as it was used to create the synthetic seismic trace.

SYNTHETIC SEISMOGRAM

From the sonic log, assuming constant density, normal incidence, no multiples or reverberations and no absorption, a synthetic seismic trace was created. The reflection coefficients for the trace were simply computed with the recursion:

$$RC_i = \frac{V_i\rho_i - V_{i-1}\rho_{i-1}}{V_i\rho_i + V_{i-1}\rho_{i-1}} \tag{1}$$

where RC_i is the reflection coefficient at the ith interface and V and ρ represent velocity and density respectively. The density was assumed to be constant. The depth-to-time conversion was simply done by integrating the vertical travel time

$$t_i = \sum_{j=1}^{i-1} \frac{2 * \Delta z_j}{V_j} \tag{2}$$

where Δz_j represents the thickness of the j-th layer. In this case the layer thickness was kept constant at 10 feet.

Once converted to time, the reflectivity series was interpolated to a constant time-sampling interval using an 8-point sinc interpolator and convolved with a Ricker wavelet of 60 Hz peak frequency. The right panel of Figure 1 shows the resulting synthetic trace with time zero corresponding to the shallowest depth in the log segment (3600 feet).

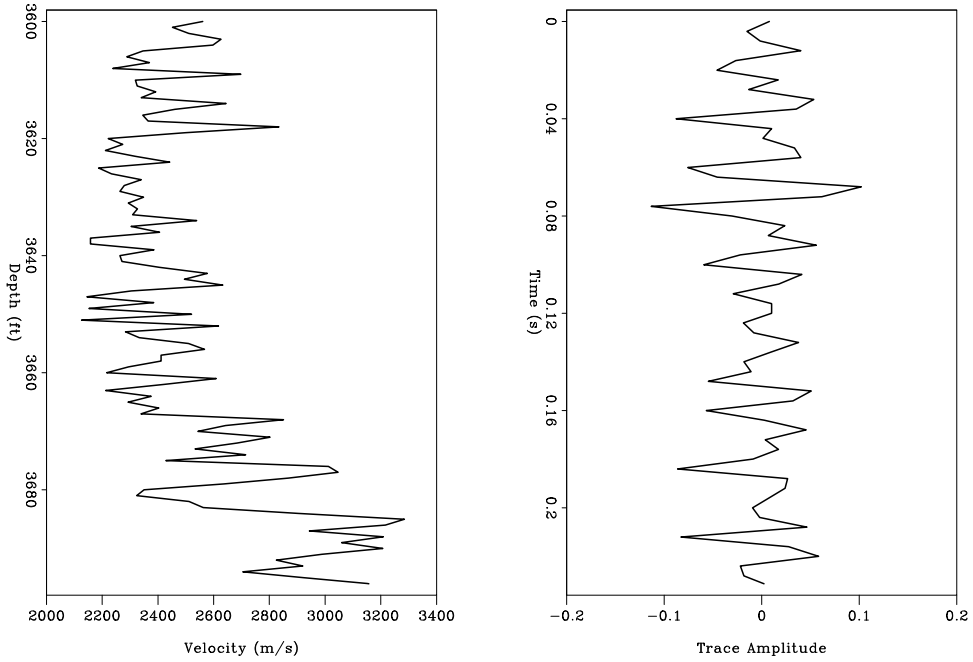


Figure 1: Left panel: Sub-sampled sonic log used to generate the synthetic trace on the right panel. The zero of the seismic trace corresponds to the initial depth of the log. gabriel2-input [ER]

VELOCITY INVERSION

The next step is to invert the seismic trace for the sonic velocities. This is done by matching candidate traces to the synthetic trace obtained from the log (right panel in Figure 1). I used a micro-genetic algorithm since it proved to be superior to a standard genetic algorithm for this problem (Alvarez, 2002). To perform the inversion, the micro-genetic algorithm starts with a small random population of synthetic sonic logs from which trial solutions, in the form of synthetic seismic traces, are generated using the procedure described in the previous section. Each of these traces is matched to the synthetic trace obtained from the real log and a figure of merit (fitness value) assigned according to the "goodness" of the match. Two possible choices of fitness functions are described below. The genetic algorithm "evolves" the solution according to the rules of evolution and survival of the fittest (Goldberg, 1989) until a satisfactory match is achieved between the synthetic seismic trace from the real log and a trial synthetic trace. Once this match is achieved, the corresponding well log will hopefully be a good match to the real one (given the restrictions of no multiple reflections and no attenuation).

This is the ideal situation but in reality we must deal with the important issue of the impossibility of directly recovering the general velocty-depth trend present in the sonic log. This happens because the seismic data lack the very low frequencies. This issue is obviously very important and will be addressed in some detail below.

Description of the Micro Genetic Algorithm Program

The micro-genetic algorithm program used in this study is generation-based (that is, children replace the parents each generation except for those parents selected by elitism), uses binary encoding with a tournament selection for choosing the parents for the next generation and uniform crossover. Other operators such as jump and creep mutation and niching were not used, following the results in (Alvarez, 2002). Table 1 summarizes the parameters of the micro-genetic algorithm.

Table 1: Summary of the micro-genetic algorithm parameters

Population size	5	Crossover rate	0.95
Mutation rate	0.0	Creep mutation rate	0.0
Minimum model parameter value	2200	Maximum model parameter value	3200
Elitism	Yes	Niching	No
Selection strategy	Tournament	Number of children	1

Parameter Encoding

Although the samples of a sonic log may not be considered statistically uncorrelated, I will assume that the samples of the sub-sampled well log are completely independent from one another. There are, therefore, 99 unknowns in our problem, one for each sample of the sub-sampled sonic log we wish to recover. Each model parameter, i.e. sample of the trial sonic log, was coded as a binary number. Although the model parameters can in principle assume any real value between some reasonable maximum and minimum values, for the purpose of representing them as binary numbers we need to discretize them. Here I used 1024 possible values for each parameter, that is, 10 bits. The maximum and minimum allowed velocities were in the range between 2200 and 3200 m/s. This means a resolution just under 1 m/s for the inversion of each parameter (each sample of the sonic log).

Fitness function

For the fitness function I tested both an L_1 and an L_2 norm of the sample-wise difference between the two traces as the criterion to measure the fit. That is, I used

$$f_j = \sum_{i=0}^{N} |x_i - y_{j,i}| \tag{3}$$

and

$$f_j = \sqrt{\sum_{i=0}^{N} (x_i - y_{j,i})^2} \tag{4}$$

where x and y are the sample amplitudes of the reference and the trial trace respectively, j represents the jth individual and N is the number of time samples. These fitness functions consider each sample in the sub-sampled well log as being completely independent of any other. Notice that I am using "fitness" here to actually mean the opposite of its usual meaning. That is, an individual will be considered to be highly fit if its fitness value, as defined above, is very low, i.e., it is a good match to the original trace.

Convergence Criteria

The ideal convergence criterion for a genetic algorithm would be one that guaranteed that each and all of the parameters converge independently (Goldberg, 1989; Beasley et al., 1993a). However, this may be too demanding or may result in too many iterations, so more relaxed convergence criteria are usually employed. Here I used four convergence criteria:

1. The fitness function value must be below a given threshold value.
2. The difference between the best and the average fitness is less than a given fraction of the fitness of the average individual.
3. The difference between the best individual of the current population and the best individual so far must be very small (even zero). This means that the most fit individual has converged even if the population itself has not.
4. the number of iterations (generations of the sample population) exceeds a given limit. This prevents the algorithm from spending too much time refining an existing solution.

The combination of these criteria is intended to guarantee that the solution is not due to a lucky guess of the random generator but to a comprehensive search of the model space.

Inversion Constraints

Constraints are critical in non-linear problems to limit the number of possible solutions and to increase the chances of finding a solution that satisfies all aspects of the problem. An obvious constraint to be imposed on the inversion is that the computed velocities be within a reasonable interval. In this case I chose $2200 < V < 3200$ in m/s. This constraint, however, is not enough because we need to account for the velocity-depth trend present in the sonic log. It is a well-known fact that in general wave propagation velocities increase with depth due to several factors such as compaction, cementation, closing of small fractures due to overburden, etc. There may be effects such as overpressure that locally change this behavior, but in average we expect a trend of velocities increasing with depth. When inverting seismic data for interval velocities, it is disappointing to find that this general trend is not recovered. The reason is that the information in the velocity trend would correspond to very low frequencies (lower than, say, 5 Hz) and these low frequencies are not recorded in the standard seismic method. Furthermore, with land data we usually attenuate or eliminate low frequencies because of the

detrimental effect of ground-roll on the seismic data. Since this trend is not resolvable by the direct inversion of the seismic trace, it is necessary to impose a second constraint to honor it. This constraint will be discussed in more detail below.

RESULTS

The top left panel of Figure 2 shows a comparison between the synthetic trace obtained from the real sonic log, and the one obtained with the genetic algorithm using the L_1 norm. The top right panel shows a similar comparison between the original and the inverted log. Although the match of the seismic traces is not perfect, the main features in the original trace were remarkably well recovered. The result for the sonic logs, however, is not so auspicious and the lack of the general velocity-depth trend makes it very difficult to evaluate the "goodness" of the match. If we forget about the trend and examine the details in the logs, we can see that most features of the log were indeed recovered, but the overall result is not satisfactory. The bottom panels show the same comparisons using the L_2 norm. The results are similar: good match of the seismic traces and poor match of the sonic logs because of our inability to recover the velocity-depth trend.

VELOCITY-DEPTH TREND

The results of the previous section illustrate the need to consider the velocity-depth trend as a constraint to the inversion. I used a Savitzky-Golay smoothing filter (Press et al., 1992) to compute the trend from the sonic log. This filter is particularly well-suited for this purpose because it is easy to control the number of samples of the input log that are used to compute a sample of the smoothed log and the degree of the smoothing polynomial.

Figure 3 shows the original sub-sampled log (top) and the computed velocity-depth trend (bottom) after applying a Savitzky-Golay filter with 33 points (16 to each side) with a sixth-order smoothing polynomial. Extrapolation was used in the original log to 16 points off each of the ends of the log to avoid end-effect problems with the filter. The velocity-depth trend is well recovered, in particular for the deeper samples. This trend will be used as a constraint to the inversion.

INVERSION WITH THE VELOCITY-DEPTH TREND

The velocity-depth trend of Figure 3 was used to constrain the possible values of velocity for each sample of the sonic log. Figure 4 shows a comparison similar to that of Figure 2 but considering the velocity-depth trend. The match for the seismic traces did not change much, since it was already very good without the trend, although a comparison of the top left panels in Figures 2 and 4 does show some improvement in the trace match (for example between times 0.04 and 0.065 s). The match of the sonic log, however, changed significantly, and now, without the distraction of the lack of the trend, we can see that the inverted sonic log matches

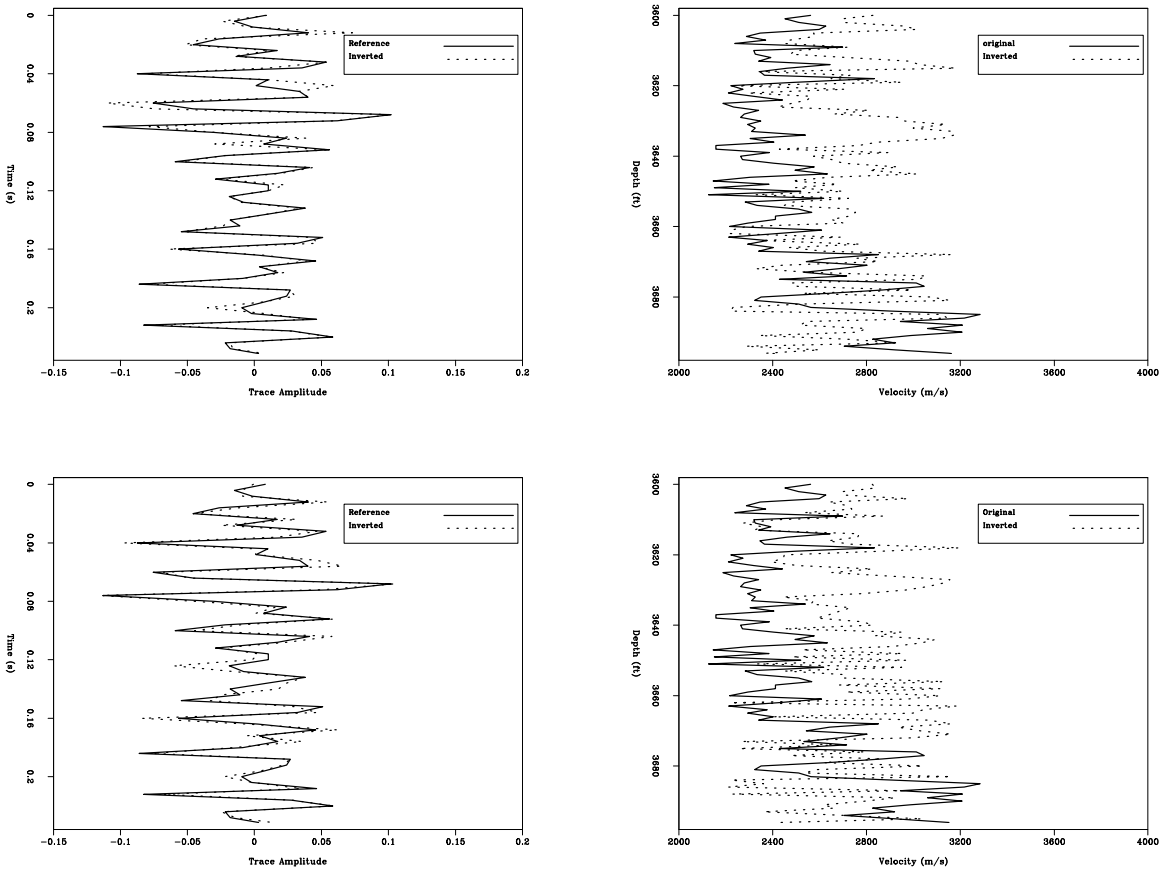


Figure 2: Comparison of the synthetic trace and the inverted trace (top left panel) and the original log and the inverted log (top right panel) using the L_1 norm. The bottom panels show similar comparisons using the L_2 norm. gabriel2-L1L2_inv_no_trend [ER]

the original log very well, except for depths between about 700 and 800 feet where the match is a little off in depth. Similar situation can be seen in the bottom panels which show the same comparison when the L_2 norm was employed. The conclusion is similar as for the L_1 norm, that is, the match of the sonic logs increases greatly, although it is still far from perfect, illustrating the limitations of the method.

Discussion

I just showed that the inclusion of the velocity-depth trend increased significantly the match between the original and the inverted sonic log. The question now is: Why didn't we get an even better match in the sonic logs? An obvious explanation is that the sonic log in this case was sub-sampled to 10 feet which is still a very thin interval when it comes to the depth resolution that we can achieve with typical seismic data. In order to get a better match we would need higher frequencies, so there is always a limit in the quality of the match that we can reasonably expect to obtain in the inversion of the sonic log. To illustrate this point,

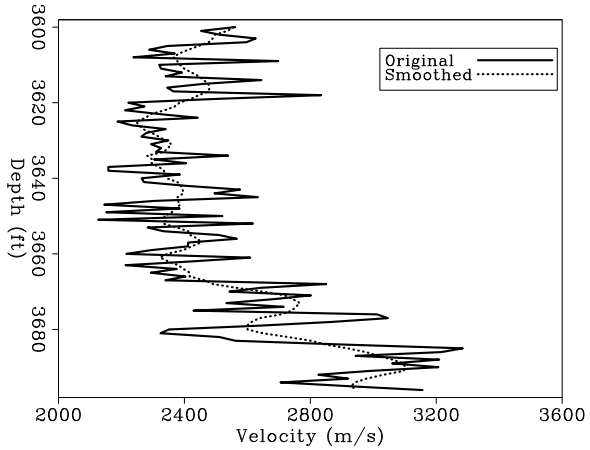


Figure 3: Sub-sampled original sonic log and velocity-depth trend computed with 33-point sixth order Savitzky-Golay filter. gabriel2-velocity_depth_trend [ER]

Figure 5 shows the same comparison as in the bottom panles of Figure 4 except that the peak frequency of the Ricker wavelet was halved. Clearly, the lack of high frequencies have a very detrimental effect in the quality of the match of the sonic logs. See in particular the depth intervals between about 300-400 ft, 450-550 ft and 700-800 ft.

It should also be noted that the example here is not very realistic because I used the exact same approach to compute the seismic data to be matched and to compute the trial solutions. In a real situation the reference trace will be obtained from a real seismic survey whereas the trial solutions will still be computed with a procedure similar to the one used here (perhaps improved by allowing random noise or multiples or attenuation). Therefore, there will be reasons other than the velocity-depth trend or the frequency content of the data that will play a role. The estimation of the seismic wavelet, for example, is well-recognized as a very important issue when inverting for real data.

CONCLUSIONS AND FUTURE WORK

I have shown with this simple example that micro-genetic algorithms are a good tool for velocity inversion of seismic data. Within the limitations of the modeling algorithm and the frequency content of the data, it is possible to get a reasonably accurate inversion of sonic log velocities provided the inversion is supplied with an estimate of the velocity-depth trend.

From the point of view of the genetic algorithm inversion some lessons have been learned after extensive parameter testing: use of a micro-genetic algorithm with uniform cross-over and no mutation emerges as the best option for this problem (as opposed to a standard genetic algorithm with single-point cross-over and jump mutation). A micro-genetic-algorithm population of 5 or 7 individuals with a cross-over probability of 0.9-0.95 seems to be optimum for this problem. The micro-genetic algorithm in this case converges to a reasonable solution after about 4000 generations (20000 function evaluations) in 10 seconds on a single-processor Linux PC.

An important issue to be further analyzed is that of the multi-modality of the search space.

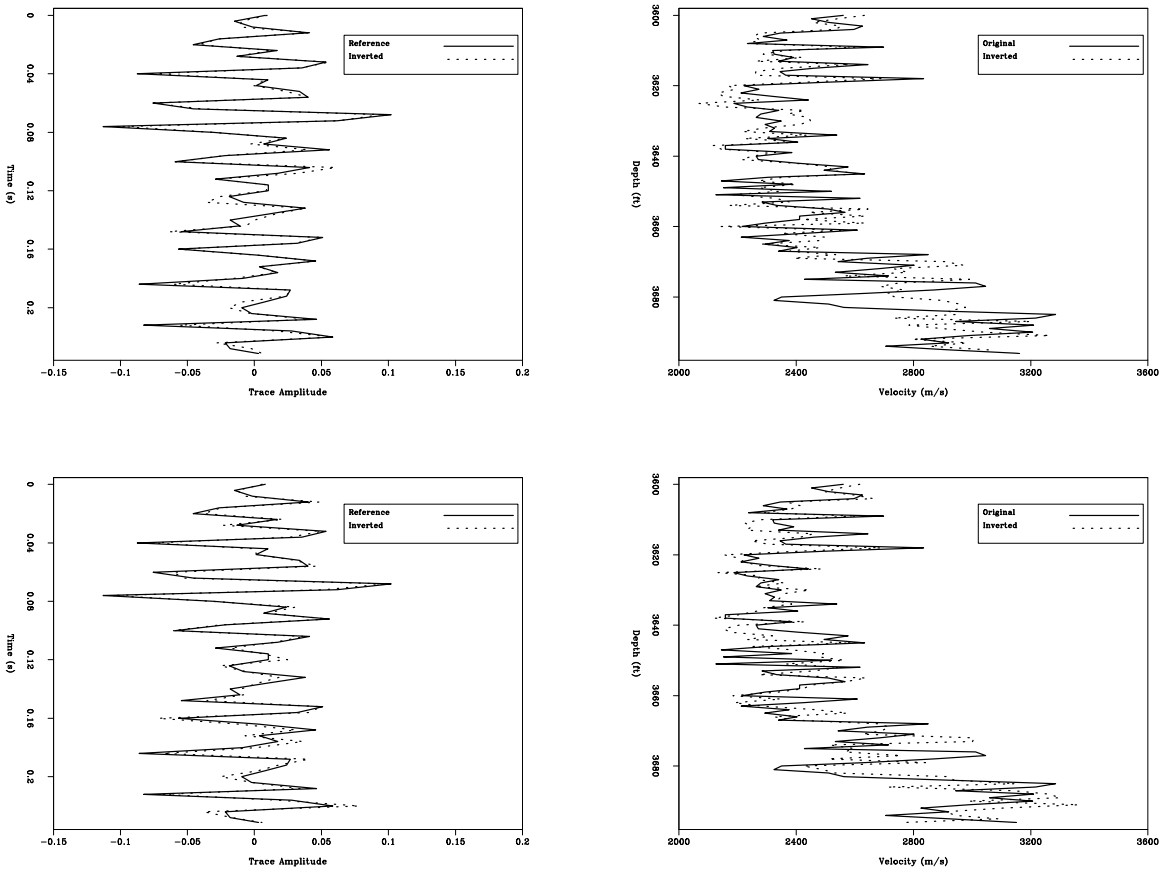


Figure 4: Results for inversion with the velocity-depth trend. Comparison of the synthetic trace and the inverted trace (top left panel) and the original log and the inverted log (top right panel) using the L_1 norm. The bottom panels show similar comparisons using the L_2 norm. gabriel2-L1L2_inv_with_trend [ER]

It is clear in this case that there is a single global minimum, namely matching the original trace sample-by-sample. This, however, does not guarantee a similar sample-to-sample match in the original log, which is a consequence of the non-linearity of the problem.

I have found that once I get close enough to this global minimum it takes a large number of iterations to scape local minima (many traces match "almost exactly" the original). My present convergence criteria do not allow for checking of convergence of individual parameters so I have to investigate alternative options. Another issue is the convenience of using the floating point representation of the model parameters directly for the inversion rather than their binary representation. This may increase the resolution of the model parameters and make the inversion overall more robust.

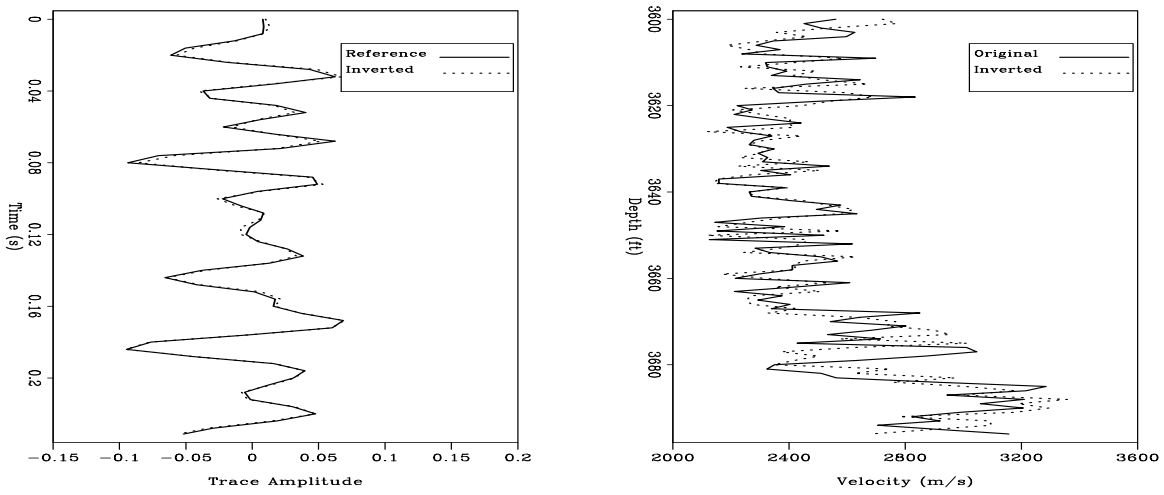


Figure 5: Comparison of reference and matched trace (left panel) and of original and inverted log using the L_2 norm when the velocity-depth trend was taken into account and the peak frequency of the Ricker wavelet was halved. gabriel2-L2_inv_trend_low_freq [ER]

REFERENCES

Alvarez, G., 2002, Can we make genetic algorithms work in high-dimensionality problems?: SEP–**112**, 195–212.

Beasley, D., Bull, D., and Martin, R., 1993a, An overview of genetic algorithms: Part 1, fundamentals: University Computing, **15**, no. 2, 58–69.

Beasley, D., Bull, D., and Martin, R., 1993b, An overview of genetic algorithms: Part 2, research topics: University Computing, **15**, no. 4, 170–181.

Falkenauer, E., 1998, Genetic algorithms and grouping problems: John Wiley and sons.

Goldberg, D., 1989, Genetic algorithms in search for optimization and machine learning: Addison-Wesley Pub. Co.

Haupt, R., and Haupt, S., 1998, Practical genetic algorithms: John Wiley and sons.

Krishnakumar, K., 1989, Micro-genetic algorithms for stationary and non-stationary function optimization.: SPIE: Intelligent control and adaptive systems, **1196**, 289–296.

Lindseth, R., 1982, Digital processing of geophysical data - areview: Soc. of Expl. Geophys.

Mavko, G., Dvorkin, J., and Mukerji, T., 2000, Rock physics handbook.: Cambridge University Press.

Press, W., Teukolsky, S., Vettering, W., and Flannery, B., 1992, Numerical recipes in fortran 77. the art of sicentific computing: Cambridge University Press.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 223–237

Modeling high-frequency acoustics velocities in patchy and partially saturated porous rock using differential effective medium theory

James G. Berryman[1]

ABSTRACT

Differential effective medium (DEM) theory is applied here to the problem of modeling physical properties of poroelastic media that are partially saturated with liquid. Typical fluid saturants are air and water, or gas and oil. If the liquid and gas saturants are homogeneously mixed, then we say the medium is partially saturated. If the liquid and gas saturants are poorly mixed, so each constituent occupies separate, but contiguous, regions of the porous medium, we say the medium has patchy saturation. Some examples are presented to show that a reasonable approach to modeling the effects of patchy saturation at high frequencies (200 kHz and above) is produced by treating the medium as if it is a composite of gas-saturated and liquid-saturated porous inclusions that are homogeneously mixed together. Estimates of the properties are obtained using differential effective medium theory. The results differ dramatically from those predicted by Gassmann's equations for homogeneous mixing of the fluids in individual pores. In particular, the shear modulus depends on the elastic properties of the fluid constituents, unlike the quasi-static behavior predicted by Gassmann.

INTRODUCTION

One of the perennial problems in rock physics has been the difficulty of understanding how seismic wave speeds in fluid saturated and partially fluid saturated rocks depend on the wave frequency. Field methods for exciting seismic waves are usually in the 1–100 Hz band, while well-logging tools might be in the 1–50 kHz band. However, the careful controlled experiments needed to verify the predictions of the theory can normally only be done at still higher frequencies, typically in the 200–1000 kHz band. This segmentation of the frequency band into its distinct regions of application has caused and continues to cause much confusion about what is known and unknown about wave propagation and attenuation in rocks.

Theoretical analyses of Gassmann (1951) and Biot (1956a,b) provide low frequency results. Gassmann's results in particular are very low frequency, really applicable to the quasi-static domain, and therefore strictly apply only to the very lowest seismic frequencies. But a skeptical scientist wants proof of these theories, and it is sometimes hard to find convincing

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

verifications among experiments done on rocks. Plona (1980) provided one very nice series of ultrasound experiments ($\simeq$ 1 MHz) using water-saturated sintered glass-bead samples (instead of rocks) showing (Chin *et al.*, 1983) that the Biot-Gassmann theory is in fact correct even in this high frequency regime, at least for such simple porous materials. The likely reason for this success is the high permeability ($1 \sim 10$D) and lack of microcracks in the porous glass-bead samples.

Difficulties still exist in explaining some high frequency laboratory data, especially in situations having low fluid permeability (and therefore making it unlikely that Gassmann's quasi-static conditions are close to being satisfied) and also having partial saturation conditions (mixtures of gas and liquid are present in the pores). There has been extensive work on partial saturation by Nur and Simmons (1969), Domenico (1974), Walls (1982), Murphy (1984), Berryman *et al.* (1988), Endres and Knight (1989), Knight and Nolen-Hoeksema (1990), Dvorkin and Nur (1998), Dvorkin *et al.* (1999a,b), Johnson (2001), and Berryman *et al.* (2002a), among others. In particular, the work of Knight and Nolen-Hoeksema (1990) makes it particularly apparent that great care must be taken in modeling these types of materials because details clearly matter. Whether the gas and liquid components are homogeneously mixed or are distributed in a patchy manner (*i.e.*, gas here, liquid there) makes a significant difference in the measured wave speeds. Even for the shear waves, where according to Gassmann's low frequency calculations we might conclude that there should be no difference at all (Berryman, 1999), we find clearly observable differences at higher frequencies.

Through a series of recent publications (Berryman *et al.*, 2000; 2002a), it has become clear that the most appropriate of the simple effective medium models for partial saturation conditions at high frequencies is the differential effective medium (DEM) theory (Berryman *et al.*, 2002b). The present paper will show specifically how to use this theory to fit data on partial and patchy saturation in a low-porosity, low-permeability granite and two tight sandstones.

DIFFERENTIAL EFFECTIVE MEDIUM THEORY

Differential effective medium (DEM) theory (Bruggeman, 1935; Cleary *et al.*, 1980; Walsh, 1980; Norris, 1985; Avellaneda, 1987) takes the point of view that a composite material may be constructed by making infinitesimal changes in an already existing composite. There are only two effective medium schemes known at present that are realizable, *i.e.*, that have a definite microgeometry associated with the modeling scheme. The differential scheme is one of these (Cleary *et al.*, 1980; Norris, 1985; Avellaneda, 1987) — and one version of the self-consistent approach (Korringa *et al.*, 1979; Berryman, 1980a,b; Milton, 1985) is the other. This fact, together with the associated analytical capabilities (including ease of computation and flexibility of application), provides strong motivation to study the predictions of both of these schemes and the differential scheme in particular. We can have confidence that the results will always satisfy physical and mathematical constraints, such as the Hashin-Shtrikman bounds (Hashin and Shtrikman, 1961; 1962).

When the inclusions are sufficiently sparse that they do not form a single connected network throughout the composite, it is most appropriate to use the Differential Effective Medium

(DEM) to model their elastic behavior (Berge *et al.*, 1993). Assume that the host material has moduli K_m and μ_m, while the inclusion material has moduli K_i and μ_i. Then, the effective bulk and shear moduli (indicated as such by the asterisks) of the composite are parametrized by $K^*(y)$ and $\mu^*(y)$ where the volume fraction of the inclusion phase is y. The equations governing the changes in these constants are then well-known to be

$$(1-y)\frac{dK^*(y)}{dy} = \left[K_i - K^*(y)\right] P^{*i} \tag{1}$$

and

$$(1-y)\frac{d\mu^*(y)}{dy} = \left[\mu_i - \mu^*(y)\right] Q^{*i}, \tag{2}$$

where the scalar factors, P^{*i} and Q^{*i}, will be explained in the following paragraph, y is porosity which equals inclusion volume fraction here, and the subscript i again stands for inclusion phase. We assume that the reader is somewhat familiar with this approach, and will therefore not dwell on its derivation, which is easily found in many places including, for example, Berryman (1992). These equations are typically integrated starting from porosity $y = 0$ with values $K^*(0) = K_m$ and $\mu^*(0) = \mu_m$, which are assumed here for modeling purposes to be the mineral moduli values for the single homogeneous solid constituent. Integration then proceeds from $y = 0$ to the desired highest value $y = \phi$ (the porosity of the sample), or possibly over the whole range to $y = 1$ for some purposes of analysis. When integrating this way, we might imagine the result is, for example, simulating cracks being introduced slowly into a granite-like solid. The same procedure can be used for a sandstone-like material assuming this medium has starting porosity $y = \phi_0$ with $K^*(\phi_0) = K_s$ and $\mu^*(\phi_0) = \mu_s$. Integration then proceeds from $y = \phi_0$ to $y = 1$. This introduction of crack (or soft) porosity into a material containing spherical (or stiff) porosity is conceptually equivalent to the porosity distribution model of Mavko and Jizba (1991).

The factors P^{*i} and Q^{*i} appearing in (1) and (2) are the so-called polarization factors for bulk and shear modulus (Eshelby, 1957; Wu, 1966). These depend in general on the bulk and shear moduli of both the inclusion, the host medium (assumed to be the existing composite medium $*$ in DEM), and on the shapes of the inclusions. The polarization factors usually have been computed from Eshelby's well-known results (Eshelby, 1957) for ellipsoids, and Wu's work (Wu, 1966) on identifying the isotropically averaged tensor based on Eshelby's formulas. These results can be found in many places including Berryman (1980b) and Mavko *et al.* (1998).

Because it is relevant both to low porosity granites and to sandstones having equant (*i.e.*, close to spherical) porosity as well as flat cracks, the case we consider here is that of penny-shaped cracks, where

$$P^{*i} = \frac{K^* + \frac{4}{3}\mu_i}{K_i + \frac{4}{3}\mu_i + \pi\alpha\gamma^*} \tag{3}$$

and

$$Q^{*i} = \frac{1}{5}\left[1 + \frac{8\mu^*}{4\mu_i + \pi\alpha(\mu^* + 2\gamma^*)} + 2\frac{K_i + \frac{2}{3}(\mu_i + \mu^*)}{K_i + \frac{4}{3}\mu_i + \pi\alpha\gamma^*}\right], \tag{4}$$

with α ($0 < \alpha < 1$) being the crack (oblate spheroidal) aspect ratio,

$$\gamma^* \equiv \mu^*[(3K^* + \mu^*)/(3K^* + 4\mu^*)], \tag{5}$$

and where the superscript $*$ identifies constants of the matrix material when the inclusion volume fraction is y. This formula is a special limit of Eshelby's results not included in Wu's paper, but apparently first obtained by Walsh (1969). Walsh's derivation assumes $\mu_i/\mu_m << 1$ and allows $K_i/K_m << 1$, with these approximations being made before any assumptions about smallness of the aspect ratio α. By taking these approximations in the opposite order, *i.e.*, letting aspect ratio be small first and then making assumptions about smallness of the inclusion constants, we would obtain instead the commonly used approximation for disks. But this latter approximation is actually quite inappropriate for the bulk modulus when the inclusion phase is a gas such as air (for then the ratio $K_i/K_m << 1$) or for the shear modulus when the inclusion phase is any fluid (for then $\mu_i \equiv 0$), as the formulas become singular in these limits. This is why the penny-shaped crack model is commonly used instead for cracked rocks.

In general the DEM equations (1) and (2) are coupled, as both equations depend on both the bulk and shear modulus of the composite. This coupling is not a serious problem for numerical integration. Later in the paper, we will show results obtained from integrating the DEM equations numerically.

HIGH FREQUENCIES

Figure 1 displays data from one granite and two sandstones in the frequency range 200–1000 kHz (also see Table 1). These examples were chosen for common display to emphasize the fact that there can be very clear deviations from the Gassmann-Domenico predictions at high frequencies. In particular, we see the startling difference in the right-hand subplots (Figure 1b,d,f) that the slopes of the patchy saturation lines (*i.e.*, lines connecting data points for fully dry and fully saturated samples) in all three cases are negative, instead of positive as predicted for low-frequency behavior [see Berryman *et al.* (2002a)]. Nevertheless, all three plots on the left (Figure 1a,c,e) seem to behave in a manner consistent with the Gassmann-Domenico ideas. Sierra White granite and Schuler-Cotton Valley sandstone (both measured at about 200 kHz) show behavior consistent with our interpretation of nearly ideal patchy saturation, again consistent with the drainage method of producing the changes in saturation.

TABLE 1. Some physical parameters of the samples considered in Figure 1. (Note: 1 mD $\simeq 1 \times 10^{-15}$ m^2.)

Sample	*Porosity (%)*	*Permeability (mD)*	*Grain Size (μm)*
Sierra White granite[a]	0.8	–	–
Schuler-Cotton Valley sandstone[a]	3.3	1.1×10^{-3}	100
Spirit River sandstone[b]	5.2	1.0×10^{-3}	125–150

[a] Murphy (1982)
[b] Knight and Nolen-Hoeksema (1990); Murphy (1982); Walls (1982)

On the other hand, the Spirit River sandstone (Figures 1e,f) was measured in the 600 kHz to 1 MHz frequency range, and two distinct methods of saturation were employed. The drainage method in this case again seems to show patchy saturation content, although it is not very close to the ideal patchy saturation line. The imbibition data are expected to produce a more uniform distribution of gas and liquid in the pores than that obtainable in most cases with a drainage method. Thus, imbibition data should behave much as predicted by Gassmann-Domenico, at least at low frequencies. Here we observe in Figure 1e that the imbibition data do indeed mimic the predicted behavior of Gassmann-Domenico, even though we are at high frequencies. Taken together, these results seem to suggest that something fairly simple is happening to produce these data, and that the main issue in Figure 1b,d,f is probably the actual violation of the Gassmann's very low frequency result that the shear modulus is not influenced by the presence of the fluid, and/or how such behavior can be modelled.

Seifert *et al.*(1999), working at about 1 MHz, chose to use the symmetric effective medium theory of Berryman (1980a) to model their data. The frequencies used are low enough so that a typical wavelength is 2 mm, while the grain sizes for the sands studied range from 210 to 250 μm, so the wavelength is an order of magnitude larger than the grain sizes and effective medium theory can safely be used. For an unconsolidated sand fully saturated with liquid, such a system is fairly closely approximated by a fluid suspension and therefore the self-consistent scheme (Berryman, 1980a) is appropriate for their problem. However, it would not be appropriate for our partial saturation problem where the pore fluid is sometimes all gas, and the solid frame always plays the major role in supporting both compressional and shear stresses. The better choice for such problems is a differential effective medium (DEM) theory [see Berge *et al.* (1995) for a more complete discussion of the advantages and disadvantages of these methods]. Then the solid can be treated correctly as the host medium and the gas and liquid constituents are treated strictly as inclusion phases — a requirement for this problem.

Our calculation for patchy saturation first uses DEM to compute the bulk and shear moduli for a porous solid saturated with gas only, and then repeats the calculation for bulk and shear moduli for a porous solid saturated with liquid only. In both cases, the shear modulus starts out at the shear modulus of the solid host medium and this is gradually replaced by (zero) inclusion shear modulus as the final desired porosity is attained. Nevertheless, the results in the gas- and liquid-saturated cases differ in these calculations because even though they have the same value for inclusion shear modulus, they do not have the same value for inclusion bulk modulus. This difference is important to the computed results. The physical reason for the difference is that in a random medium when a shear stress is applied macroscopically, it is resolved microscopically into both shear and compressional component stresses [see Berryman and Wang (2001) for an analysis of this aspect of the problem]. Trapped liquid can support some of those resolved compressional stresses resulting from an applied shear stress and therefore makes the saturated porous medium stronger in both shear and compression than when the same medium is saturated instead with a gas. Thus, the theory shows that $\mu_{dr} \neq \mu_{sat}$ when the saturating fluid is a liquid. This result disagrees with Gassmann, but does not contradict Gassmann. The point is that Gassmann's result is quasi-static and therefore pertinent for much lower frequencies, wherein the fluid can respond to the applied shear field by simply moving out of the way. But for trapped fluids or relatively rapid wave propagation through the medium,

the result just described must hold.

At high enough frequencies, adding liquid to a partially saturated system will in fact increase the effective shear modulus of the system. Thus, when we plot λ/μ versus ρ/μ, it is no longer the case that ρ/μ is a monotonic function of saturation. The density ρ is still a monotonic increasing function of saturation S as before, but now μ is also a monotonic increasing function of saturation. Therefore, the ratio ρ/μ is not necessarily monotonic and its behavior depends on which of the competing changes in the numerator and denominator dominate. The results for Sierra White granite (very low porosity) in Figure 1b clearly show that the main effect of addition of liquid to the system is to produce changes in μ at low porosities, with the result that the patchy saturation line has the opposite sign of slope (seen in Figure 1b) as that predicted by Gassmann-Domenico (seen in Figure 1a) and the data tend to fall along this line. The results are similar but not quite so well behaved for the Schuler-Cotton Valley sandstone in Figure 1d.

The most interesting behavior is observed for the Spirit River sandstone in Figure 1f. Here we see very clearly that as the liquid saturation increases, at first we have an increase in ρ/μ and then, when some special value of saturation (near 40% for the drainage data) is achieved, the influence of liquid on the shear modulus becomes more important and dominates the remainder of the curves up to full saturation.

Of the examples shown here, all three deviate dramatically from the predicted Gassmann-Domenico behavior. All these cases have the lowest porosities and permeabilities of the examples considered by Berryman *et al.* (2002a). This effect is presumably related therefore to the influence of permeability on the inability of the pore fluid pressure to equilibrate during the passage time of the wave, *i.e.*, having a higher likelihood of acoustically disconnected porosity.

EXAMPLES

When modeling rocks using effective medium theory, or really any approach, it is important to minimize the number of choices (size and shape of the inclusions, etc.) available to the modeler. Especially when dealing with cracks and simultaneously with partial and patchy saturation, the number of possible scenarios multiplies rapidly. For example, it would be entirely realistic to assume that there is a distribution of both crack aspect ratios and sizes present in the rocks. But since this distribution is surely not known, we will assume instead that there is only a single aspect ratio of crack present and choose values to lie in the range $\alpha = 0.001$–0.1. The decision to have only one aspect ratio is arbitrary. But it is motivated by the need to minimize the nonuniqueness inherent in the enterprise of fitting these data. Occam's razor applies here: we try to use the simplest possible model (a single aspect ratio), and if we cannot fit the data then we have learned something about the rock. We will find however that the simplest model is always sufficient when fitting the velocity data alone. The range of values that are considered sensible are based in part on the data of Hadley (1976) on Westerly granite, where it was observed that α ranged from 10^{-4} all the way up to unity, with a mode around 10^{-3}.

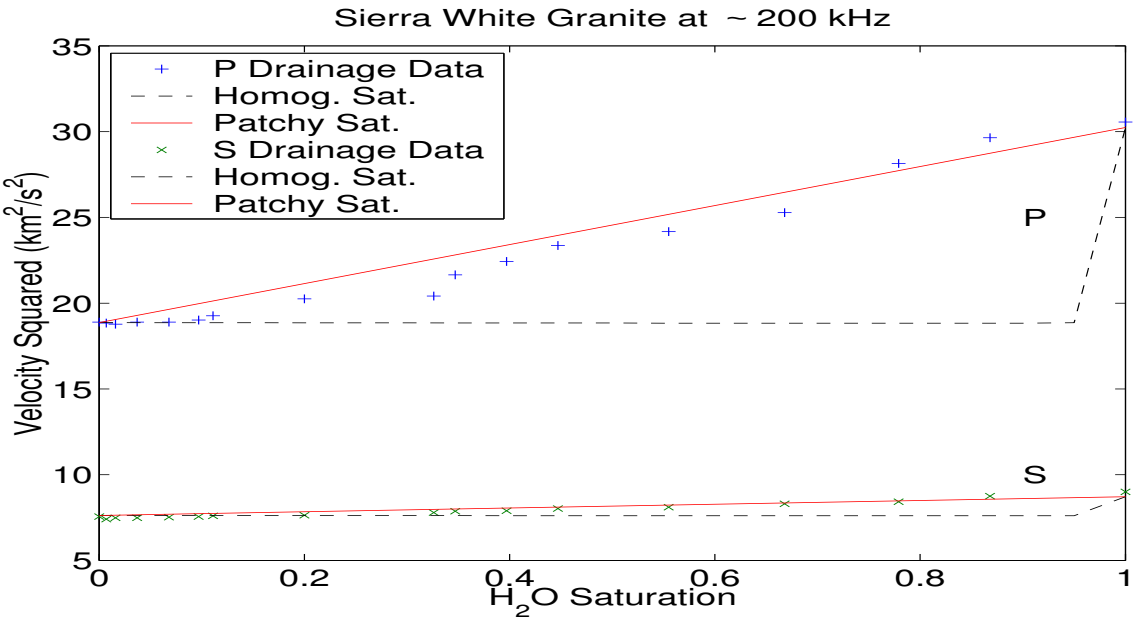


Figure 1: Square of the velocity data for the Sierra White granite measured by Murphy (1982) at 200 kHz. The dashed lines are the DEM results for compressional and shear when it is assumed that the saturation is homogeneous in each pore. The solid lines are the results for patchy saturation. Clearly, the data all fall closer to the patchy saturation lines at the higher values of liquid saturation. For the very lowest values of liquid saturation, the data seem to mimic the homogeneous saturation curve. jim1-swgvsq_all2 [NR]

Sierra White granite

All the data presented here for Sierra White granite are from Murphy (1982). This case is especially simple as the porosity is quite low ($\phi = 0.008$) and therefore the effect of liquid saturation on the density is very small ($< 0.3\%$ density effect), which we will treat for purposes of hand analysis as negligible. Thus, essentially the entire effect of liquid saturation depends on how the liquid is distributed in the pores and how this affects the bulk and shear moduli only. We will model this simply by considering the effects of gas saturation and liquid saturation separately and then combining the results [Voigt (1928) average] for the patchy saturation effects. For homogeneous saturation, we use DEM with an effective fluid bulk modulus given by the Reuss (1929) average (harmonic mean) of the fluids' moduli.

Murphy describes Sierra White as a granite "composed of a sparse population of low aspect ratio cracks, embedded in a composite of elastic grains." Some preliminary calculations done for the present work indicate that an aspect ratio of $\alpha \simeq 0.005$ to 0.02 should give results very consistent with the measured values for Sierra White. To fit the data at both the fully gas saturated end and the fully water saturated end, we found that $K_m = 57.7$ GPa and $\mu_m = 31.7$ GPa were good choices. The computation was performed at 21 equally spaced values of saturation for homogeneous saturation. The patchy saturation curve is obtained by connecting the two end points on a plot of velocity squared with a straight line. (For situations with

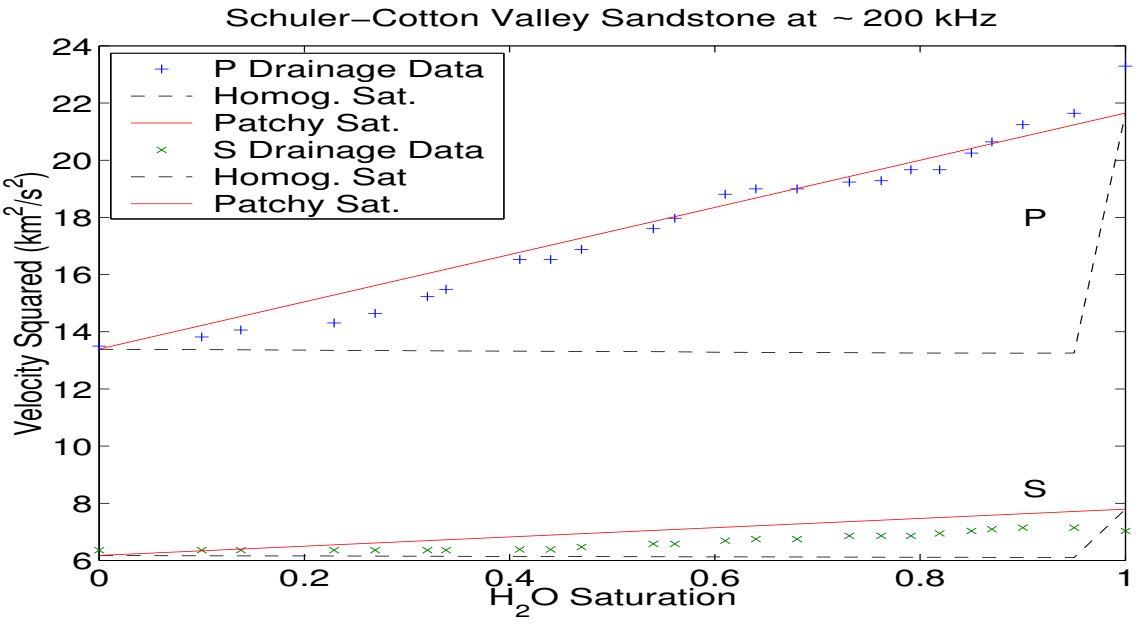


Figure 2: Square of the velocity data for the Schuler-Cotton Valley sandstone measured by Murphy (1982) at about 200 kHz. The dashed lines are the DEM results for compressional and shear when it is assumed that the saturation is homogeneous in each pore. The solid lines are the results for patchy saturation. In this case, all the data for compressional waves fall closer to the patchy saturation line than to the curve for homogeneous saturation. The data for shear waves seem to be much less sensitive to the assumed crack aspect ratio than were the compressional wave data. The results seem to suggest that the cracks that dominate compressional wave propagation are not the same as the ones dominating shear wave propagation for this sample. jim1-scvvsq_all2 [NR]

significant density variation, it is preferable to plot the Lamé constants λ and μ instead of the squares of the velocities — but for small porosity this is always a small difference that we choose to ignore here.) The results are shown in Figure 2. Clearly, all the data fall closer to the patchy saturation lines at the higher values of liquid saturation. For the very lowest values of liquid saturation, the data seem to mimic the homogeneous saturation curve, but at these low saturation levels the two curves are very close together anyway.

Schuler-Cotton Valley sandstone

The data presented here on Schuler-Cotton Valley sandstone are from Murphy (1982). Walls (1982) also studied permeability variations of other samples from the same formation, but both of his samples had higher porosity than the one studied by Murphy ($\phi = 0.033$). Murphy's compressional wave measurements were made at about 200 kHz and the shear measurements at about 150 kHz. The fitting parameters chosen for this sample were $K_m = 41.8$ GPa, $\mu_m = 36.7$ GPa, and $\alpha = 0.015$. Otherwise the analysis was identical to that for Sierra White granite. Results are shown in Figure 3. These curves are similar to those for the granite since the

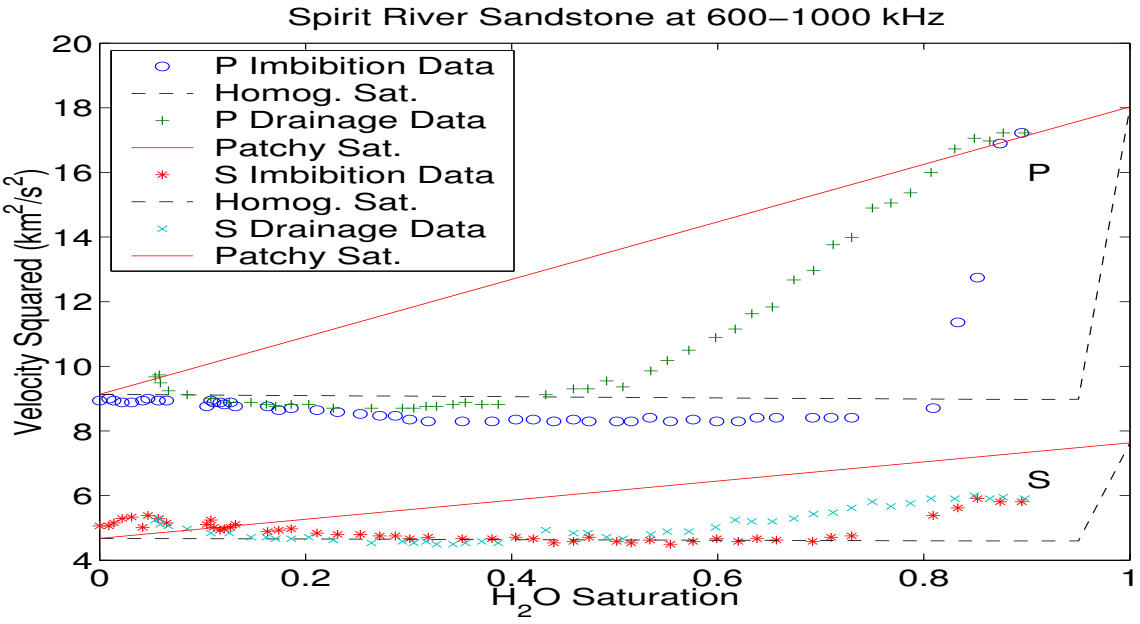


Figure 3: Square of the velocity data for the Spirit River sandstone measured by Knight and Nolen-Hoeksema (1990) in the range 600-1000 kHz. The dashed lines are the DEM results for compressional and shear when it is assumed that the saturation is homogeneous in each pore. The solid lines are the results for patchy saturation. Imbibition data are expected to fall closer to the homogeneous saturation curve, which is seen here for both V_p and V_s. Although the drainage data are expected to fall closer to the patchy saturation line, this tendency is only observed here at the higher saturation levels. See discussion in the text. jim1-srsvsq_all2 [NR]

compressional wave data show a very clear trend along the patchy saturation line. The shear wave data were harder to fit for this case, since the change in shear wave speed does not mimic that of the compressional wave as one might expect. For the Sierra White granite, a single value of α was sufficient to bring both P- and S-wave speeds into good agreement with the theoretical curves. Here, this was not possible, as the variation in compressional wave speed is substantially greater than that for the shear wave speed. A possible conclusion from this observed behavior is that the cracks that dominate compressional wave speed changes are not the same as those for the shear wave speed changes in this rock. But if this is true, it is not clear how to go about modeling such an effect.

Spirit River sandstone

The acoustical data on Spirit River sandstone are from Knight and Nolen-Hoeksema (1990). Both Murphy (1982) and Walls (1982) also studied the Spirit River sandstone, and Walls' sample SR6547 is apparently the same one studied by Knight and Nolen-Hoeksema. Imbibition data are expected to fall closer to the homogeneous saturation curve, which is seen to be true here for both V_p and V_s. Although the drainage data are expected to fall closer to the patchy

saturation line, this tendency is only observed here at the higher saturation levels (> 40 %).

Walls (1982) studied the gas permeability of the Spirit River sandstone as a function of both saturation and effective pressure. At room pressure the gas permeability changed from 100 μD to 47 μD as the liquid saturation changed from zero to 40%. Then, as the effective pressure increased, the permeability of the 40% saturated sample fell more rapidly than that of the other sample, differing by an order of magnitude at 30 MPa.

Since the porosity of this sample is about 5%, the volume fraction of the whole sample occupied by liquid at 40% saturation is about 2%. It is commonly observed that liquids can begin to percolate (*i.e.*, maintain a continuous connected pathway) across a porous sample when their volume fraction is of the order of 2-3%. So we assume that this dramatic departure of the observed drainage data is in some way related to this percolation threshold.

TABLE 2. Fitting parameters of the three rock samples considered in the text.

Sample	K_m (GPa)	μ_m (GPa)	α
Sierra White granite	57.7	31.7	0.005
Schuler-Cotton Valley sandstone	41.8	36.7	0.015
Spirit River sandstone	30.0	69.0	0.0125

Summary of results

The fitting parameters used to match the data in all three of the examples shown are listed in Table 2. For comparison, the values of bulk and shear moduli for quartz are often quoted as $K_m = 37.0$ GPa and $\mu_m = 44.0$ GPa, respectively. The fitting parameters obtained here lie in the ranges $30.0 \leq K_m \leq 57.7$ GPa and $31.7 \leq \mu_m \leq 69.0$ GPa. According to Walls (1982), the mineralogy of some other Schuler-Cotton Valley sandstones had about 72% quartz with the next most common mineral being quartz overgrowth ranging from 7-12%. The specific sample of Spirit River sandstone used by both Walls (192) and Knight and Nolen-Hoeksema (1990) had only about 34% quartz and 30% siderite, with the next most abundant components being chert, dolomite, and quartz overgrowth – each being in the range 7-10%. Siderite and dolomite both have significantly different, stronger bulk and shear moduli compared to quartz. So we conclude that the range of values observed in our fitting parameters are all quite possible, physically reasonable values but hard to check otherwise.

The observed grain sizes for Schuler-Cotton Valley sandstone and Spirit River sandstone were 100 μm and 125–150 μm, respectively. At the frequencies used in the experiments, the wavelengths for Schuler-Cotton Valley were $\lambda_p \simeq$ 18–24mm, and $\lambda_s \simeq$ 18mm. For Spirit River, the wavelengths were $\lambda_p \simeq$ 3–4mm and $\lambda_s \simeq$ 4mm. So, for Schuler-Cotton Valley, the grain size over wavelength is about 0.005, while, for Spirit River, it was about 0.035. We expect that the effective medium theory approach should be valid whenever these ratios are less than about 0.3, and certainly for an order of magnitude or more as is the case for these samples.

CONCLUSIONS

The main point of this exercise has been to show that the DEM equations are appropriate to use in this context (since the wavelengths are sufficiently long compared to the grain sizes) and that the DEM equations do in fact predict that right kinds of behavior in the high frequency range ($>$ 200 kHz). Gassmann's equations have clearly failed in this region (*i.e.*, the shear modulus is not independent of the fluid content), as would be expected. The quasi-static assumptions explicitly used in Gassmann's derivation are not satisfied in this frequency regime, and particularly so in rocks having low permeabilities (μD), as is the case for all the samples considered here. We can understand both qualitatively and semi-quantitatively what is happening in these samples by making use of DEM as a modeling tool. To do more detailed modeling requires much more detailed information about the constituents, their spatial distribution, their bonding characteristics, and the distribution and character of voids and cracks. We are still some ways from being able to determine all of these parameters in real rocks, but nevertheless can conclude that the methods described and used for modeling here do correctly capture the physics of these complicated high-frequency acoustics problems.

ACKNOWLEDGMENTS

I thank Bill Murphy and Rosemary Knight for providing access to their unpublished data files. I thank Pat Berge and Brian Bonner for helpful suggestions, and Morgan Brown for helpful comments that improved the manuscript.

REFERENCES

Avellaneda, M., 1987, Iterated homogenization, differential effective medium theory and applications: *Commun. Pure Appl. Math* **40**, 527–554.

Berge, P. A., Bonner, B. P., and Berryman, J. G., 1995, Ultrasonic velocity-porosity relationships for sandstone analogs made from fused glass beads: *Geophysics* **60**, 108–119.

Berryman, J. G., 1980a, Long-wavelength propagation in composite elastic media. I. Spherical inclusions: *J. Acoust. Soc. Am.* **68**, 1809–1819.

Berryman, J. G., 1980b, Long-wavelength propagation in composite elastic media II. Ellipsoidal inclusions: *J. Acoust. Soc. Am.* **68**, 1820–1831.

Berryman, J. G., 1992, Single-scattering approximations for coefficients in Biot's equations of poroelasticity: *J. Acoust. Soc. Am.* **91**, 551–571.

Berryman, J. G., 1999, Origin of Gassmann's equations: *Geophysics* **64**, 1627–1629.

Berryman, J. G., Berge, P. A., and Bonner, B. P., 2000, Transformation of seismic velocity data to extract porosity and saturation values for rocks: *J. Acoust. Soc. Am.* **107**, 3018–3027.

Berryman, J. G., Berge, P. A., and Bonner, B. P., 2002a, Estimating rock porosity and fluid saturation using only seismic velocities: *Geophysics* **67**, 391–404.

Berryman, J. G., Pride, S. R., and Wang, H. F., 2002b, A differential scheme for elastic properties of rocks with dry or saturated cracks: *Geophys. J. Int.*, accepted for publication in 2002.

Berryman, J. G., Thigpen, L., and Chin, R. C. Y., 1988, Bulk elastic wave propagation in partially saturated porous solids: *J. Acoust. Soc. Am.* **84**, 360–373.

Berryman, J. G, and Wang, H. F., 2001, Dispersion in poroelastic systems: *Phys. Rev. E* **64**, 011303-1–011303-16.

Biot, M. A., 1956a, Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range: *J. Acoust. Soc. Am.* **28**, 168–178.

Biot, M. A., 1956b, Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range: *J. Acoust. Soc. Am.* **28**, 179–1791.

Biot, M. A., 1962, Mechanics of deformation and acoustic propagation in porous media: *J. Appl. Phys.* **33**, 1482–1498.

Bruggeman, D. A. G., 1935, Berechnung verschiedener physikalischer Konstanten von heterogenen Substanzen: *Ann. Physik. (Leipzig)* **24**, 636–679.

Chin, R. C. Y., Berryman, J. G., and Hedstrom, G. W., 1985, Generalized ray expansion for pulse propagation and attenuation in fluid-saturated porous media: *Wave Motion* **7**, 43–66.

Cleary, M. P., Chen, I.-W., and Lee, S.-M., 1980, Self-consistent techniques for heterogeneous media: *ASCE J. Eng. Mech.* **106**, 861–887.

Domenico, S. N., 1974, Effect of water saturation on seismic reflectivity of sand reservoirs encased in shale: *Geophysics* **39**, 759–769.

Dvorkin, J., Mavko, G., and Nur, A., 1999a, Overpressure detection from compressional- and shear-wave data: *Geophys. Res. Lett.* **26**, 3417–3420.

Dvorkin, J., Moos, D., Packwood, J. L., and Nur, A. M., 1999b, Identifying patchy saturation from well logs: *Geophysics* **64**, 1756–1759.

Dvorkin, J. and Nur, A., 1998, Acoustic signatures of patchy saturation: *Int. J. Solids Struct.* **35**, 4803–4810.

Eshelby, J. D., 1957, The determination of the elastic field of an ellipsoidal inclusion, and related problems: *Proc. Roy. Soc. London A* **241**, 376–396.

Gassmann, F., 1951, Über die elastizität poröser medien: *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, **96**, 1–23.

Hadley, K., 1976, Comparison of calculated and observed crack densities and seismic velocities in Westerly granite: *J. Geophys. Res.* **81**, 3484–3494.

Hashin, Z., and Shtrikman, S., 1961, Note on a variational approach to the theory of composite elastic materials: *J. Franklin Inst.* **271**, 336–341.

Hashin, Z., and Shtrikman, S., 1962, A variational approach to the theory of elastic behaviour of polycrystals: *J. Mech. Phys. Solids* **10**, 343–352.

Johnson, D. L., 2001, Theory of frequency dependent acoustics in patchy-saturated porous media: *J. Acoust. Soc. Am.* **110**, 682-694.

Knight, R., and Nolen-Hoeksema, R., 1990, A laboratory study of the dependence of elastic wave velocities on pore scale fluid distribution: *Geophys. Res. Lett.* **17**, 1529–1532.

Korringa, J., Brown, R. J. S., Thompson, D. D., and Runge, R. J., 1979, Self-consistent imbedding and the ellipsoidal model for porous rocks: *J. Geophys. Res.* **84**, 5591–5598 (1979).

Mavko, G., and Jizba, D., 1991, Estimating grain-scale fluid effects on velocity dispersion in rocks: *Geophysics* **56**, 1940–1949.

Mavko, G., Mukerji, T., and Dvorkin, J., 1998, *The Rock Physics Handbook: Tools for Seismic Analysis in Porous Media*, Cambridge University Press, Cambridge, pp. 307–309.

Milton, G. W., 1985, The coherent potential approximation is a realizable effective medium scheme: *Comm. Math. Phys.* **99**, 463–500.

Murphy, William F., III, 1982, *Effects of Microstructure and Pore Fluids on the Acoustic Properties of Granular Sedimentary Materials*, Ph.D. Dissertation, Stanford University.

Norris, A. N., 1985, A differential scheme for the effective moduli of composites: *Mech. Mater.* **4**, 1–16.

Nur, A., and Simmons, G., 1969, The effect of saturation on velocity in low porosity rocks: *Earth and Planet. Sci. Lett.* **7**, 183–193.

Plona, T. J., 1980, Observation of a second bulk compressional wave in a porous medium at ultrasonic frequencies: *Appl. Phys. Lett.* **36**, 259–261.

Reuss, A., 1929, Berechung der Fleissgrenze von Mischkristallen: *Z. Angew. Math. Mech.* **9**, 55.

Seifert, P. K., B. Kaelin, and L. R. Johnson, 1999, Effect on ultrasonic signals of viscous pore fluids in unconsolidated sand: *J. Acoust. Soc. Am.* **106**, 3089–3094.

Voigt, W., 1928, *Lehrbuch der Kristallphysik*, Teubner, Leipzig, p. 962.

Walls, J. D., 1982, Tight gas sands – permeability, pore structure: and clay: *J. Petroleum Technology* **34**, 2708–2714.

Walsh, J. B., 1965, The effect of cracks on the compressibility of rock: *J. Geophys. Res.* **70**, 381–389.

Walsh, J. B., 1980, Static deformation of rock: *ASCE J. Engng. Mech.* **106**, 1005–1019.

Wu, T. T., 1966, The effect of inclusion shape on the elastic moduli of a two-phase material: *Int. J. Solids Struct.* **2**, 1–8.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 237–259

Dispersion of waves in porous cylinders with patchy saturation

James G. Berryman,[1] *Steven R. Pride*[2]

ABSTRACT

Laboratory experiments on wave propagation through saturated and partially saturated porous media have often been conducted on porous cylinders that were initially fully saturated and then allowed to dry while continuing to acquire data on the wave behavior. Since it is known that drying typically progresses from outside to inside, a sensible physical model of this process is concentric cylinders having different saturation levels — the simplest example being a fully dry outer cylindrical shell together with a fully wet inner cylinder. We use this model to formulate the equations for wave dispersion in porous cylinders for patchy saturation (*i.e.*, drainage) conditions. In addition to multiple modes of propagation obtained numerically from these dispersion relations, we find two distinct analytical expressions for torsional wave modes. We solve the dispersion relation for torsional waves for two examples: Massillon sandstone and Sierra White granite. The drainage analysis appears to give improved agreement with the data for both these materials.

INTRODUCTION

The classic work of Pochhammer (1876) and Chree (1886) gave exact solutions for wave propagation in elastic rods. When the rod is instead a porous cylinder with fluid-filled pores, the equations of linear elasticity do not describe all possible motions of the fluid/porous-solid mixture. Biot's theory of fluid-saturated porous media provides a continuum theory, permitting the fluid and solid components to move independently and accounts approximately for the attenuation of waves due to viscous friction. Gardner (1962) used Biot's theory (Biot, 1956a,b) to study long-wavelength extensional waves in circular cylinders. Gardner considered only the low-frequency regime where the second bulk compressional mode predicted by Biot's theory is diffusive in character. Gardner also limited consideration to the case of open-pore surface boundary conditions.

The present work is based in part on another paper by Berryman (1983), in which both open-pore and closed-pore surface boundary conditions for the fluid-saturated porous cylinder were studied. Here we consider only the open-pore surface, but we allow non-uniform or patchy saturation (Berryman, 1988; Knight and Nolen-Hoeksema, 1990; Knight *et al.*, 1998; Johnson, 2001) inside the cylinder. In particular, it is quite common to study partial satura-

[1]**email**: berryman@sep.stanford.edu
[2]**email**: spride@univ-rennes1.fr

tion in the laboratory under drainage or drying conditions wherein an initially fully saturated porous cylinder is allowed to dry while continuing to acquire data on the cylinder's modes of oscillation. We want to model this behavior explicitly. The simplest such model is concentric cylinders with a fully dry outer cylindrical shell enclosing a fully liquid-saturated inner cylinder. A more realistic model would involve many layers with various degrees of partial saturation between the dry outer shell and the saturated inner cylinder, but such complications will not be treated here. We find that studies of the two-layer case have all the important physical complications expected in this problem, while still having enough simplicity that some of the analysis can be done semi-analytically — thereby providing soughtafter insight into the problem.

We present the equations of poroelasticity, and then show the forms of the equations needed for cylindrical geometry. Appropriate boundary conditions for our problem are discussed. Equations are subsequently formulated to determine both the extensional and torsional modes of concentric poroelastic cylinders under conditions of partial saturation. Solutions of these equations are computed and discussed here for torsional waves, while the harder problem of extensional waves will be treated fully in a later publication.

EQUATIONS OF POROELASTICITY

For long-wavelength disturbances ($\lambda >> h$, where h is a typical pore size) propagating through a porous medium, we define average values of the (local) displacements in the solid and also in the saturating fluid. The average displacement vector for the solid frame is $\mathbf{u}$ while that for the pore fluid is $\mathbf{u}_f$. The average displacement of the fluid relative to the frame is $\mathbf{w} = \phi(\mathbf{u} - \mathbf{u}_f)$. For small strains, the frame dilatation is

$$e = e_x + e_y + e_z = \nabla \cdot \mathbf{u}, \tag{1}$$

where e_x,e_y,e_z are the Cartesian strain components. Similarly, the average fluid dilatation is

$$e_f = \nabla \cdot \mathbf{u}_f \tag{2}$$

(e_f also includes flow terms as well as dilatation) and the increment of fluid content is defined by

$$\zeta = -\nabla \cdot \mathbf{w} = \phi(e - e_f). \tag{3}$$

With these definitions, Biot (1962) obtains the stress-strain relations in the form

$$\delta\tau_{xx} = He - 2\mu(e_y + e_z) - C\zeta, \tag{4}$$

and similarly (with permutations) for the other compressional components $\delta\tau_{yy}$,$\delta\tau_{zz}$, while

$$\delta\tau_{zx} = \mu\left(\frac{\partial u_x}{\partial z} + \frac{\partial u_z}{\partial x}\right), \tag{5}$$

and again for $\delta\tau_{yz}$,$\delta\tau_{xy}$ for the other shear components. And finally, for the fluid pressure,

$$\delta p_f = M\zeta - Ce. \tag{6}$$

The $\delta\tau_{ij}$ are deviations from equilibrium of average Cartesian stresses in the saturated porous material and δp_f is similarly the isotropic pressure deviation in the pore fluid.

With time dependence of the form $\exp(-i\omega t)$, the coupled wave equations that incorporate (4)-(6) are of the form

$$\begin{aligned} \omega^2(\rho\mathbf{u} + \rho_f\mathbf{w}) &= C\nabla\zeta - (H-\mu)\nabla e - \mu\nabla^2\mathbf{u}, \\ \omega^2(\rho_f\mathbf{u} + q\mathbf{w}) &= M\nabla\zeta - C\nabla e, \end{aligned} \tag{7}$$

where $\rho = \phi\rho_f + (1-\phi)\rho_m$ is the bulk-density of the material and $q = \rho_f\left[\alpha/\beta + i\,F(\xi)\eta/\kappa\omega\right]$ is the effective density of the fluid in relative motion. The kinematic viscosity of the liquid is η; the permeability of the porous frame is κ; the dynamic viscosity factor is given, for our choice of sign for the frequency dependence, by $F(\xi) = \frac{1}{4}\{\xi T(\xi)/[1 + 2T(\xi)/i\xi]\}$, where $T(\xi) = \frac{\mathrm{ber}'(\xi) - i\,\mathrm{bei}'(\xi)}{\mathrm{ber}'(\xi) - i\,\mathrm{bei}'(\xi)}$ and $\xi = (\omega h^2/\eta)^{\frac{1}{2}}$. The functions ber($\xi$) and bei($\xi$) are the real and imaginary parts of the Kelvin function. The dynamic parameter h is a characteristic length generally associated with and comparable in magnitude to the steady-flow hydraulic radius. The tortuosity $\alpha \geq 1$ is a pure number related to the frame inertia which has been measured by Johnson *et al.* (1982) and has also been estimated theoretically by Berryman (1980a,b).

The coefficients H, C, and M are given by [see Gassmann (1951), Geertsma (1957), Biot and Willis (1957), Geertsma and Smit (1961), and Stoll (1974)]

$$H = K + \frac{4}{3}\mu + (K_m - K)^2/(D-K), \tag{8}$$

$$C = K_m(K_m - K)/(D-K), \tag{9}$$

and

$$M = K_m^2/(D-K), \tag{10}$$

where

$$D = K_m[1 + \phi(K_m/K_f - 1)]. \tag{11}$$

Equations (8)-(11) are correct as long as the porous material may be considered homogeneous on the microscopic scale as well as the macroscopic scale.

To decouple the wave equations (7) into Helmholtz equations for the three modes of propagation, we note that the displacements $\mathbf{u}$ and $\mathbf{w}$ can be decomposed as

$$\mathbf{u} = \nabla\Upsilon + \nabla\times\boldsymbol{\beta}, \qquad \mathbf{w} = \nabla\psi + \nabla\times\boldsymbol{\chi}, \tag{12}$$

where Υ, ψ are scalar potentials and $\boldsymbol{\beta}$, $\boldsymbol{\chi}$ are vector potentials. Substituting (12) into (7), we find (7) is satisfied if two pairs of equations are satisfied:

$$(\nabla^2 + k_s^2)\boldsymbol{\beta} = 0, \qquad \boldsymbol{\chi} = -\rho_f\boldsymbol{\beta}/q \tag{13}$$

and

$$(\nabla^2 + k_\pm^2)A_\pm = 0. \tag{14}$$

The wavenumbers in (13) and (14) are defined by

$$k_s^2 = \omega^2(\rho - \rho_f^2/q)/\mu \tag{15}$$

and

$$k_\pm^2 = \tfrac{1}{2}\left\{b + f \mp [(b-f)^2 + 4cd]^{\frac{1}{2}}\right\}, \tag{16}$$

$$\begin{aligned} b &= \omega^2(\rho M - \rho_f C)/\Delta,\ c = \omega^2(\rho_f M - qC)/\Delta, \\ d &= \omega^2(\rho_f H - \rho C)/\Delta,\ f = \omega^2(qH - \rho_f C)/\Delta, \end{aligned} \tag{17}$$

with $\Delta = MH - C^2$. The linear combination of scalar potentials has been chosen to be $A_\pm = \Gamma_\pm \Upsilon + \psi$, where

$$\Gamma_\pm = d/(k_\pm^2 - b) = (k_\pm^2 - f)/c. \tag{18}$$

With this identification (18) of the coefficients $\Gamma_\pm$, the decoupling is complete.

Equations (13) and (14) are valid for any choice of coordinate system. They may be applied to boundary value problems with arbitrary symmetry. Biot's theory will therefore be applied to porous elastic cylinders in the next section.

EQUATIONS FOR A POROUS CYLINDER

To work most easily in cylindrical geometry, we rewrite the stress-strain relations (4)-(6) in cylindrical coordinates. If z is the coordinate along the cylinder axis while r and θ are the radial and azimuthal coordinates, it is not difficult to show that

$$\delta\tau_{rr} = He - 2\mu(e_\theta + e_z) - C\zeta, \tag{19}$$

$$\delta\tau_{r\theta} = \mu\left(\frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} + \frac{1}{r}\frac{\partial u_r}{\partial \theta}\right), \tag{20}$$

$$\delta\tau_{rz} = \mu\left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r}\right), \tag{21}$$

and (6) for δp_f remains unchanged. The stress increments $\delta\tau_{zz}$, $\delta\tau_{\theta\theta}$, and $\delta\tau_{\theta z}$ are not of direct interest in the present application. The dilatations are given by

$$e = e_r + e_\theta + e_z, \tag{22}$$

where

$$e_r = \frac{\partial u_r}{\partial r}, \quad e_\theta = \frac{u_r}{r} + \frac{1}{r}\frac{\partial u_\theta}{\partial \theta}, \quad e_z = \frac{\partial u_z}{\partial z}. \tag{23}$$

We redefine potential $\boldsymbol{\beta}$ in terms of two scalar potentials according to

$$\boldsymbol{\beta} = \hat{\mathbf{z}}\beta_1 + \nabla \times (\hat{\mathbf{z}}\beta_2), \tag{24}$$

where both β_i satisfy

$$(\nabla^2 + k_s^2)\beta_i = 0 \qquad \text{for} \qquad i = 1, 2. \tag{25}$$

For the problem of interest here, we will have two distinct regions: The first region is a cylinder centered at the origin, within which solutions of (14) and (25) must be finite at the origin. Results take the form

$$A_\pm = \alpha_\pm J_0(j_\pm) \exp i(k_z z - \omega t), \tag{26}$$
$$\beta_1 = \gamma_s J_0(j_s) \exp i(k_z z - \omega t), \tag{27}$$
$$\beta_2 = (\alpha_s / i k_z) J_0(j_s) \exp i(k_z z - \omega t), \tag{28}$$

where

$$j_\pm = k_{\pm r} r, \qquad j_s = k_{sr} r \tag{29}$$

and

$$k_{\pm r}^2 = k_\pm^2 - k_z^2, \qquad k_{sr}^2 = k_s^2 - k_z^2. \tag{30}$$

J_0 is the zero-order Bessel function of the first kind. The coefficients $\alpha_\pm$, α_s, γ_s, are constants to be determined from the boundary conditions.

The second region is a cylindrical shell around the first region. In this region, the factors $k_\pm$ and k_s take different values from the those in the central region, indicated by $k_\pm^*$ and k_s^* (where $*$ means air-filled, and does *not* ever mean complex conjugate in this paper). Furthermore, two linearly independent solutions of the equations are allowed, *i.e.*, both J_0 and Y_0 (the Bessel function of the second kind, sometimes known as the Neumann function). In the outer shell, we have four coefficients apiece for J_0 and Y_0, all of which must also be determined by the boundary conditions.

Noting that

$$\Upsilon = (A_+ - A_-)/(\Gamma_+ - \Gamma_-),$$
$$\psi = (A_+\Gamma_- - A_-\Gamma_+)/(\Gamma_- - \Gamma_+) \tag{31}$$

from the definitions of $A_\pm$, and substituting (26)-(28) and (31) into (12), and the result into (6) and (19)-(21), we finally obtain

$$\delta\tau_{r\theta} = m_{11}\gamma_s \equiv -\mu k_{sr}^2 J_2(j_s)\gamma_s, \tag{32}$$
$$\delta\tau_{rr} = a_{11}\alpha_+ + a_{12}\alpha_- + a_{13}\alpha_s, \tag{33}$$
$$-\delta p_f = a_{21}\alpha_+ + a_{22}\alpha_- + a_{23}\alpha_s, \tag{34}$$
$$\delta\tau_{rz} = a_{31}\alpha_+ + a_{32}\alpha_- + a_{33}\alpha_s, \tag{35}$$

where

$$(\Gamma_+ - \Gamma_-)a_{11} = \left[(C\Gamma_- - H)k_+^2 + 2\mu k_z^2\right] J_0(j_+) + 2\mu k_{+r} J_1(j_+)/r, \tag{36}$$

$$(\Gamma_+ - \Gamma_-)a_{12} = -2\mu k_{-r}^2 J_1(j_-)/j_- + \left[(H - C\Gamma_+)k_-^2 - 2\mu k_z^2\right] J_0(j_-), \tag{37}$$

$$a_{13} = 2\mu k_{sr}^2 J_2(j_s), \tag{38}$$

$$(\Gamma_+ - \Gamma_-)a_{21} = (M\Gamma_- - C)k_+^2 J_0(j_+), \tag{39}$$

$$(\Gamma_+ - \Gamma_-)a_{22} = (C - M\Gamma_+)k_-^2 J_0(j_-), \tag{40}$$

$$(\Gamma_+ - \Gamma_-)a_{31} = -2i\mu k_z k_{+r} J_1(j_+), \tag{41}$$

$$(\Gamma_+ - \Gamma_-)a_{32} = 2i\mu k_z k_{-r} J_1(j_-), \tag{42}$$

$$ik_z a_{33} = -\mu(k_s^2 - 2k_z^2)k_{sr} J_1(j_s), \tag{43}$$

and $a_{23} = 0$. There is an implicit factor of $\exp i(k_z z - \omega t)$ on the right-hand side of (32)-(35).

Berryman (1983) has shown that a_{11}, a_{13}, a_{31}, and a_{33} reduce in the limit $\phi \to 0$ to the corresponding results for isotropic elastic cylinders by Pochhammer (1876), Chree (1886, 1889), Love (1941), and Bancroft (1941), as they should.

BOUNDARY CONDITIONS

Appropriate boundary conditions for use with Biot's equations have been considered by Deresiewicz and Skalak (1963), Berryman and Thigpen (1985), and Pride and Haartsen (1996) and we make use of these results here.

At the external surface $r = R_2$ where the outer porous material contacts the surrounding air, it is appropriate to use the free surface conditions

$$-\delta p_f = 0,\ \delta\tau_{rr} = 0,\ \delta\tau_{r\theta} = 0,\ \delta\tau_{rz} = 0, \tag{44}$$

for the deviations from static equilibrium. If the cylinder is sealed on $r = R_2$, then the first of these needs to be replaced by $w_r = 0$.

The internal interface at $r = R_1$ needs more precise definition. We assume that all the meniscii that are separating the inner fluid from the outer fluid are contained within a thin layer

(shell) of thickness δh (a few grain sizes in width) straddling the surface $r = R_1$. All fluid that enters this interface layer goes into stretching the meniscii since as Pride and Flekkoy (1999) have shown, it is reasonable to assume that the contact lines of the meniscii remain pinned under seismic stressing. The locally incompressible flow conserves fluid volume so that the rate at which the inner fluid enters the interface layer is equal to the rate at which the outer fluid leaves the layer thus requiring

$$\dot{w}_r(r = R_1 + \delta h/2) = \dot{w}_r^*(r = R_1 - \delta h/2). \tag{45}$$

This and the following conditions are to be understood in the limit where $\delta h/R_1 \to 0$. It is also straightforward to obtain the standard results

$$\tau_{rr} = \tau_{rr}^*, \;\; \tau_{r\theta} = \tau_{r\theta}^*, \;\; \tau_{rz} = \tau_{rz}^*, \tag{46}$$

and

$$\dot{u}_r = \dot{u}_r^*, \;\; \dot{u}_\theta = \dot{u}_\theta^*, \;\; \dot{u}_z = \dot{u}_z^*. \tag{47}$$

The final condition to establish on $r = R_1$ is that involving the fluid pressure.

The rate at which energy fluxes radially through the porous material is given by $\tau_{ri}\dot{u}_i - p_f\dot{w}_r$ with implicit summation over the index i. The difference in the rate at which energy is entering and leaving the interface layer is due to work performed in stretching the meniscii. Each meniscus has an initial mean curvature H_o that is determined by the initial fluid pressures (those that hold before the wave arrives) as $p_{f0} - p_{f0}^* = \sigma H_o$ where σ is the surface tension. As the wave passes, the ratio between the actual mean curvature H and H_o is a small quantity on the order of the capillary number $\epsilon = \eta|\dot{w}_r|/\sigma$ [see Pride and Flekkoy (1999)] where $|\dot{w}_r|$ is some estimate of the induced Darcy flow and that goes as wave strain times wave velocity ($|\dot{w}_r| < 10^{-3}$ m/s). Since $\sigma > 10^{-2}$ Pa·m for air-water interfaces, we have $\epsilon < 10^{-4}$, which can be considered negligible. By integrating the energy flux rate over a Gaussian shell that straddles $r = R_1$, it is straightforward to obtain

$$\begin{aligned} &[\tau_{ri}\dot{u}_i - (p_{f0} + \delta p_f)\dot{w}_r] - \\ &[\tau_{ri}^*\dot{u}_i^* - (p_{f0}^* + \delta p_f^*)\dot{w}_r^*] = \sigma H_o \dot{w}_r[1 + O(\epsilon)]. \end{aligned} \tag{48}$$

Thus, since all components here except fluid pressure are continuous, we find that, when ϵ is small compared to unity,

$$\delta p_f = \delta p_f^*. \tag{49}$$

In other words, to the extent that the capillary number can be considered small (always the case for linear wave problems), the wave-induced increments in fluid pressure are continuous at $r = R_1$.

To apply the boundary conditions (45) and (49), we need in addition to (34) the result

$$w_r = a_{41}\alpha_+ + a_{42}\alpha_- + a_{43}\alpha_s, \tag{50}$$

where

$$(\Gamma_+ - \Gamma_-)a_{41} = k_{+r} J_1(j_+)\Gamma_-, \tag{51}$$

$$(\Gamma_+ - \Gamma_-)a_{42} = -k_{-r} J_1(j_-)\Gamma_+, \tag{52}$$

$$a_{43} = k_{sr} J_1(j_s)\rho_f / q. \tag{53}$$

The remaining stress conditions (46) are determined by (33) and (35).

To apply the boundary conditions (47), we need the explicit expressions for the displacement which follow from (12). The results are of the form

$$u_r = a_{51}\alpha_+ + a_{52}\alpha_- + a_{53}\alpha_s, \tag{54}$$

where

$$(\Gamma_+ - \Gamma_-)a_{51} = -k_{+r} J_1(j_+), \tag{55}$$

$$(\Gamma_+ - \Gamma_-)a_{52} = k_{-r} J_1(j_-), \tag{56}$$

$$a_{53} = k_{sr} J_1(j_s), \tag{57}$$

and

$$u_z = a_{61}\alpha_+ + a_{62}\alpha_- + a_{63}\alpha_s, \tag{58}$$

where $a_{61} = a_{62} = 0$, and

$$a_{63} = k_{sr}^2 J_0(j_s)/ik_z. \tag{59}$$

Both (54) and (58) are needed for extensional waves, while the remaining component,

$$u_\theta = m_{21}\gamma_s \equiv k_{sr} J_1(j_s)\gamma_s, \tag{60}$$

is needed only for torsional waves. As before, there is an implicit factor of $\exp i(k_z z - \omega t)$ on the right-hand side of (51)-(53), (55)-(57), and (59).

It follows from (32)-(35), (50), and (60) that γ_s (for the inner cylinder) and the corresponding coefficients for the cylindrical shell are all completely independent of the other mode coefficients and, therefore, relevant to the study of torsional waves, but not for extensional waves. Pertinent equations for the torsional wave dispersion relation are continuity of the angular displacement, u_θ, and stress, $\tau_{r\theta}$, at the internal interface, and vanishing of the stress, $\tau_{r\theta}$, at the external surface.

The final set of equations for the extensional wave dispersion relation involves nine equations with nine unknowns. The nine unknowns are: α_+, α_-, α_s (coefficients of J_0 in the central cylinder), plus three α^*'s (coefficients of J_0) and three η^*'s (coefficients of Y_0) for region of the cylindrical shell. The nine equations are: the continuity of radial and one tangential stress as well as radial and one tangential displacement at the interfacial boundary (totaling four conditions), continuity of fluid pressure and normal fluid increments across the same boundary (two conditions), and finally the vanishing of the external fluid pressure, radial and one tangential stress at the free surface (three conditions). The extensional wave dispersion relation is then determined as in Berryman (1983) by those conditions on the wavenumber k_z that result in vanishing of the determinant of the coefficients of this 9×9 complex matrix.

ELEMENTARY TORSIONAL MODES

The torsional mode of cylinder oscillation (which is trivial for a simple cylinder, porous or not) is determined here by a 3×3 system, of which 8 elements are in general nonzero. This system is therefore similar in size and difficulty to the cases studied earlier by Berryman (1983) for extensional waves in a simple fully saturated poroelastic cylinder. On the other hand, for extensional waves, the matrix determining the extensional wave dispersion relation for patchy saturation has 81 elements, of which 69 will in general be nonzero. This problem requires sufficiently different treatment from that for the torsional case that we set it aside to be studied fully in a future publication.

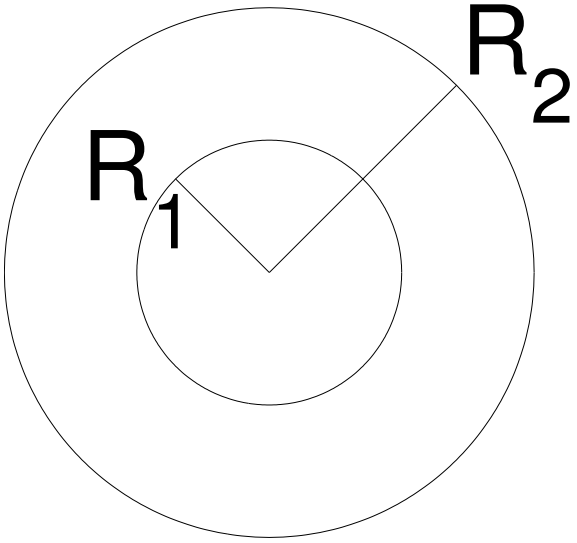


Figure 1: Cross-section of a circular cylinder, where $R_1 = S^{\frac{1}{2}} R_2$ is determined by the liquid saturation level S. jim2-concentric [NR]

We assume that the cylinder has liquid saturation level $S = (R_1/R_2)^2$, where R_2 is the radius of the cylinder and $r = R_1$ is the location of the liquid-gas interface (see Fig. 1). The dispersion relation for torsional waves is then given by

$$\begin{vmatrix} m_{11}^*(R_2) & n_{11}^*(R_2) & 0 \\ -m_{11}^*(R_1) & -n_{11}^*(R_1) & m_{11}(R_1) \\ -m_{21}^*(R_1) & -n_{21}^*(R_1) & m_{21}(R_1) \end{vmatrix} = 0, \tag{61}$$

where m_{11} and m_{21} are given by (32) and (60). The coefficients m_{11}^* and m_{12}^* have the same functional forms as m_{11} and m_{21}, but the constants are those for the shell, rather than the inner cylinder. Similarly, n_{11}^* and n_{12}^* are just the same as m_{11}^* and m_{12}^* except that J_0 and J_1 are replaced everywhere by Y_0 and Y_1, respectively.

Now we notice immediately that there could be two elementary solutions of (61), one with $m_{11}^*(R_2) = n_{11}^*(R_2) = 0$ (exterior condition) and another with $m_{11}(R_1) = m_{21}(R_1) = 0$ (interior condition). First, the interior condition is satisfied, for example, when $k_{sr} = 0$ or, equivalently, when $k_z^2 = k_s^2$. This corresponds to a torsional mode of propagation having wave speed and attenuation determined exactly by the bulk shear wave in the interior region, but the interior region is not moving since $k_{sr} = 0$ also implies that $u_\theta = 0$ from (60). Thus, the interior

condition results in the drained outer shell twisting around a stationary inner liquid-saturated cylinder. Second, the exterior condition is similarly satisfied when $k_{sr}^* = 0$ or, equivalently, when $k_z^2 = (k_s^*)^2$. This condition looks at first glance as if it might be spurious because $k_{sr}^* = 0$ suggests that u_θ at the exterior boundary might vanish identically, and then this would correspond to a trivial solution of the equations. However, looking closer, this is not the case, because at the external boundary

$$u_\theta = k_{sr}^* \left[J_1(j_s^*)\gamma_s^* + Y_1(j_s^*)\epsilon_s^* \right], \tag{62}$$

so as $k_{sr}^* \to 0$, the first term on the right hand side of (62) does vanish, both because $k_{sr}^* \to 0$ and also because $J_1(j_s^*) \to 0$. But the second term does not vanish in this limit because $|Y_1(j_s^*)| \to 2/\pi k_{sr}^* R_2 \to \infty$ as $k_{sr}^* \to 0$, and the product gives the finite result: $2/\pi R_2$. So this condition is not spurious, and corresponds to a torsional wave propagating with the speed and attenuation of the bulk shear wave speed in the drained shell material.

Can both of these elementary modes be excited? If we assume for the moment that Gassmann's equations (1951) [also see Berryman (1999)] apply to the sample, then $\mu^* = \mu$ and the only changes in shear wave velocity in the two regions are those induced by the changes in mass. In this situation, the wave speed in the air/gas saturated region will be faster than that in the water/liquid saturated region, since liquid is more dense than gas. Thus, the real part of k_s^* is smaller than that of k_s, and while the condition $(k_{sr}^*)^2 = 0$ implies that the real part of k_{sr}^2 is positive, the condition $k_{sr}^2 = 0$ implies that the real part of $(k_{sr}^*)^2$ is negative. Therefore, assuming (as we generally do here) that the attenuation in the system is relatively small, the condition $k_z = k_s^*$ leads to a propagating wave, while $k_z = k_s$ leads to a strongly evanescent wave. Note that, if Gassmann's results do not apply to the system (say at ultrasonic frequencies), then the results of the preceeding paragraph may need to be reconsidered. In particular, if the shear modulus changes rapidly with the introduction of liquid saturant, it is possible that the shear wave speed for a liquid saturated porous material may be higher than that for the gas saturated case. In this situation, all the inequalities of the preceding paragraph would be reversed, and then the condition $k_z = k_s$ leads to a propagating wave, while $k_z = k_s^*$ leads to a strongly evanescent wave.

Our conclusion then is that both modes can indeed be excited, but probably not simultaneously in the same system in the same frequency band. In a highly dispersive porous system and with broadband acoustic signal input, it could happen that both modes are propagating simultaneously in time, but in distinct/disjoint frequency bands.

HIGHER ORDER TORSIONAL MODES

For fully saturated porous cylinders, the factor that determines the torsional modes of propagation is $m_{11}(r)$ in (32). The critical factor here is the Bessel function $J_2(j_s)$ and, specifically, the whereabouts of its zeroes. One source of this information, to five figure accuracy, is the reference of Abramowitz and Stegun (1965), which provides not only the location of the zeroes $j_{2,n}$, but also the values of the corresponding derivatives $J_2'(j_{2,n})$. Having these derivatives is useful for improving the accuracy of the zeroes with a Newton-Raphson iterative method,

based on $j_{2,n} = j_{2,n}^{old} - J_2(j_{2,n}^{old})/J_2'(j_{2,n}^{old})$. This approach gives a very rapid improvement to the values of the $j_{2,n}$'s within 2 to 3 iterations. The results to order $n = 3$ are shown in TABLE 1.

TABLE 1. The zeroes $j_{2,n}$ of $J_2(z)$ as a function of the order n of appearance along the real axis.

Order n	$j_{2,n}$
0	0.00000000000000
1	5.13562230184068
2	8.41724414039987
3	11.61984117214906

Having already understood the zeroth order contributions to the dispersion relation (61) due to zeroes of k_{sr} and k_{sr}^*, we are now free to consider that neither of these factors vanishes for the higher order modes. This assumption permits us to factor these wavenumbers in or out of the determinant whenever it is convenient to do so. In particular, we note that the first two columns of (61) would have a common factor of $\mu^*(k_{sr}^*)^2$ (which could then be safely eliminated) if we first multiply the bottom row by a factor of $\mu^* k_{sr}^*$. Having made these simplifications, we find

$$\begin{vmatrix} J_2(k_{sr}^* R_2) & Y_2(k_{sr}^* R_2) & 0 \\ J_2(k_{sr}^* R_1) & Y_2(k_{sr}^* R_1) & \mu k_{sr} J_2(k_{sr} R_1) \\ J_1(k_{sr}^* R_1) & Y_1(k_{sr}^* R_1) & \mu^* k_{sr}^* J_1(k_{sr} R_1) \end{vmatrix} = 0, \tag{63}$$

after also eliminating a common factor of -1 from the top row, and $-k_{sr}$ from the third column.

Expanding the determinant along the third column, we have

$$\begin{aligned} 0 = {} & \mu^* k_{sr}^* J_1(k_{sr} R_1) \begin{vmatrix} J_2(k_{sr}^* R_2) & Y_2(k_{sr}^* R_2) \\ J_2(k_{sr}^* R_1) & Y_2(k_{sr}^* R_1) \end{vmatrix} \\ & - \mu k_{sr} J_2(k_{sr} R_1) \begin{vmatrix} J_2(k_{sr}^* R_2) & Y_2(k_{sr}^* R_2) \\ J_1(k_{sr}^* R_1) & Y_1(k_{sr}^* R_1) \end{vmatrix}. \end{aligned} \tag{64}$$

Some elementary consequences of this equation are: (a) As $R_1 \to 0$ so there is no liquid left in the system, $J_1(k_{sr} R_1)$ and $J_2(k_{sr} R_1) \to 0$ like R_1, while $Y_1(k_{sr}^* R_1)$ and $Y_2(k_{sr}^* R_1) \to \infty$ like $1/R_1$. So the dispersion relation is always satisfied in the limit when $J_2(k_{sr}^* R_2) = 0$, which is exactly the condition for the fully dry cylinder as expected. (b) If $R_1 \to R_2$, then the first determinant vanishes identically. The second determinant does not vanish in general since it approaches the Wronskian $J_2 Y_1 - J_1 Y_2 = 2/\pi k_{sr}^* R_2$, so the condition becomes $k_{sr} J_2(k_{sr} R_1) = 0$, again as expected. (c) The special case of $k_{sr} \to 0$ does not affect these conclusions, as both $J_1(k_{sr} R_1)$ and $J_2(k_{sr} R_1) \to 0$ in this limit, as they should. (d) The only case that is missing

from (64) is the one for $k^*_{sr} \to 0$. But this multiple zero of the original dispersion relation (61) was eliminated when we removed two factors of $(k^*_{sr})^2$ from the first and second column in the first step of our simplification of the dispersion relation – a step which is always legitimate except when $k^*_{sr} \equiv 0$.

We conclude that, with the one trivial exception just noted, these simplifications have kept the basic nature of the dispersion relation intact.

Lower frequency results

At lower frequencies in the range $f < 1kHz$, we may typically expect that Gassmann's results hold for the poroelastic medium, where $\mu^* \equiv \mu$. Also, to a very good approximation $k^*_s \simeq k_s$, where the only deviations from equality are those due to the differences in the densities of liquid and gas constituents. So deviations from this approxiation are most substantial when the porosity is high. From (63), we see that if the the products μk_{sr} and $\mu^* k^*_{sr}$ are equal, then these factors can be removed from the third column of the determinant. Then, the resulting third column can be subtracted from the first column, and the result can be expanded along the first column to give:

$$\frac{2J_2(k_{sr}R_2)}{\pi k_{sr}R_1} = 0, \tag{65}$$

having again used the fact that $J_2(z)Y_1(z) - J_1(z)Y_2(z) = 2/\pi z$. So the important zeroes in this case are again those of J_2, some of which are already displayed in TABLE 1.

Ignoring the imaginary part of k, which is usually quite small in the limit, we have the analytical result that

$$v_z^{(n)} = \frac{v_s}{[1-(j_{2,n}v_s/\omega R_2)^2]^{1/2}}. \tag{66}$$

Thus, at the higher frequencies, this velocity approaches that of the shear wave as expected. When the lower frequencies are approached, there is an obvious cutoff frequency, $f_c^n = j_{2,n}v_s/2\pi R_2$, below which these torsional modes do not propagate for $n \geq 1$. Since this low frequency cutoff may often be in conflict with the assumption under consideration here (*i.e.*, frequencies low enough that Gassmann's equation is satisfied), we expect generally that very few of the higher order modes can be excited in this limit. The main result is therefore that $v_z = v_z^0 = v_s$ is the velocity that will be observed in laboratory experiments in this frequency domain, with only very few exceptions.

We will not consider this rather special case any further in this paper.

Higher frequency results

The more interesting case is that for higher frequencies, in which case it is now understood (Berryman *et al.*, 2000, 2002; Berryman and Wang, 2001) that the simple Biot-Gassmann theory is actually inadequate because there can be dependence of μ on liquid saturant properties at high frequencies, such as $f > 1kHz$. The precise frequency at which this becomes

important is material dependent, but it is generally observed that for ultrasonic frequencies $f > 20kHz$ some deviations from Biot-Gassmann predictions are normally found. So it is in this regime that the distinctions between air-filled and water-filled pores become important for the torsional motion of a cylinder.

Case: $\mu^* = \mu$

Even if there is no difference between μ^* and μ, there can still be significant differences between k_s^* and k_s due to the differences in the fluid viscosities and densities of liquid and gas constituents. So we will treat this case next.

For Massillon sandstone, Murphy (1982, 1984) measured extensional and shear wave velocities at $f = 560$ Hz over a range of partial saturations produced using the drainage method. Relevant properties of this sandstone are listed in TABLE 2.

TABLE 2. Properties of Massillon sandstone used in Murphy's experiments (Murphy, 1982; 1984) and Spirit River sandstone in Knight and Nolen-Hoeksema's experiments (Knight and Nolen-Hoeksema, 1990).

Property	Massillon	Spirit River
Porosity (%)	23.0	5.2
Permeability (mD)	7.37×10^2	1.0×10^{-3}
Grain size (μm)	150–200	125–150

Case: $\mu^* \leq \mu$

The presence of liquid in the pores may alter the mechanical behavior of rocks under shear deformations in at least two quite distinct ways: (a) It is often observed that a very small amount of some liquids can cause chemical interactions that tend to soften the binding material present among the grains of such a system. When this happens, the shear modulus is usually observed to decrease. (See for example FIG. 3 for Sierra White granite.) So this situation implies that $\mu^* \geq \mu$, contrary to Gassmann's results. Although this situation is well-known in practice, we will ignore it in our modeling efforts. Our justification for this will be that the medium we are calling "dry" should in fact be termed "drained" in the sense that it has been wetted previously and therefore has these chemical softening effects already factored into the modulus μ^*. In any case, our goal is not to fit data for specific rocks, but rather to understand general trends. (b) The other situation that can also occur in practice – particularly at higher frequencies – is that the liquid saturating the porous material can have a nonnegligible mechanical effect (Berryman and Wang, 2001) that tends to strengthen the medium under shear loading so that $\mu \geq \mu^*$. If this strengthening effect is great enough (and there are experimental results (see

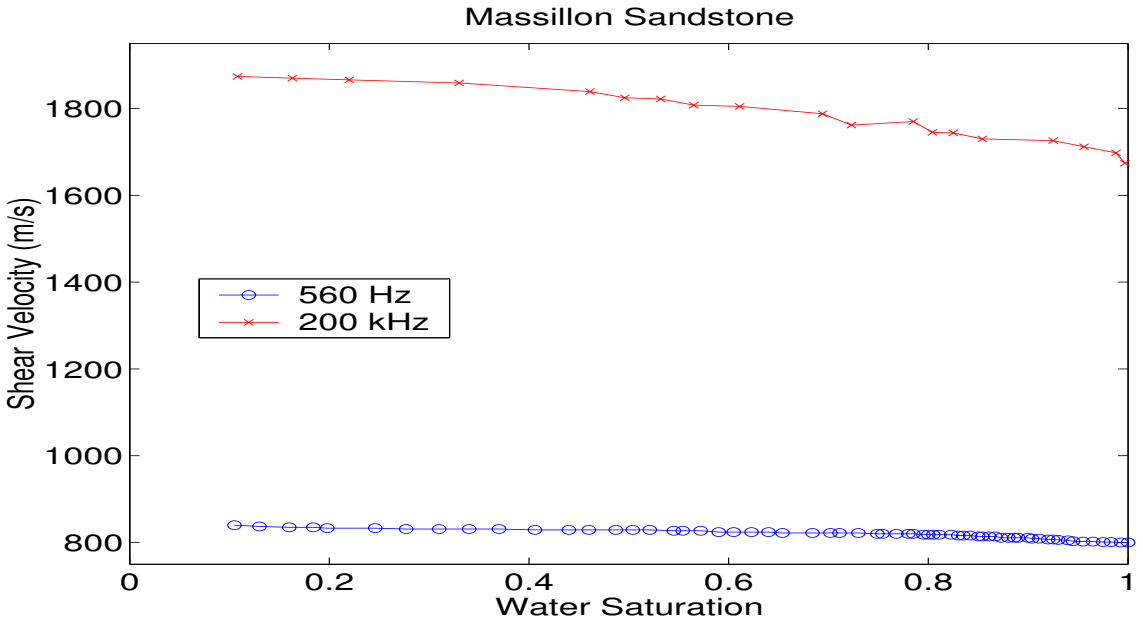


Figure 2: Shear wave velocities as a function of water saturation for drainage experiments of Murphy (1982,1984). jim2-massillon_vs [NR]

FIG. 4) that confirm this does happen in practice (Berryman *et al.*, 2002), then it is possible the density effect is more than counterbalanced by the enhanced shear modulus effect with the result that the speed of shear wave propagation in the liquid saturated medium is greater than that in the air saturated case. Depending on details of the liquid distribution in the pores, either of these cases can be included in the analysis that we now pursue in this subsection.

For Massillon sandstone, Murphy (1982,1984) also measured extensional and shear wave velocities at $f = 200$ kHz over a range of partial saturations produced using the drainage method. Relevant properties of this sandstone were listed before in TABLE 2.

SOLVING THE DISPERSION RELATION

Solving the full complex dispersion relation is somewhat tedious, and we will not try to explain this in detail here. Instead we will show results for two cases: first the Massillon sandstone (at 560 Hz) and then the Sierra White granite (at about 200 kHz). We might expect based just on the frequencies that the sandstone behavior will be close to that predicted by Gassmann, while that of the granite may differ from Gassmann.

An important observation concerning how to proceed with the analysis follows from the fact that we are seeking a curve in the complex plane, points along the curve depending on the level of saturation S. We know (at least in principle) the locations of the end points of this curve since they are exactly the points for full liquid saturation and full gas saturation. If we assume that the attenuation is relatively small so the wavenumbers k_s and k_s^* have small

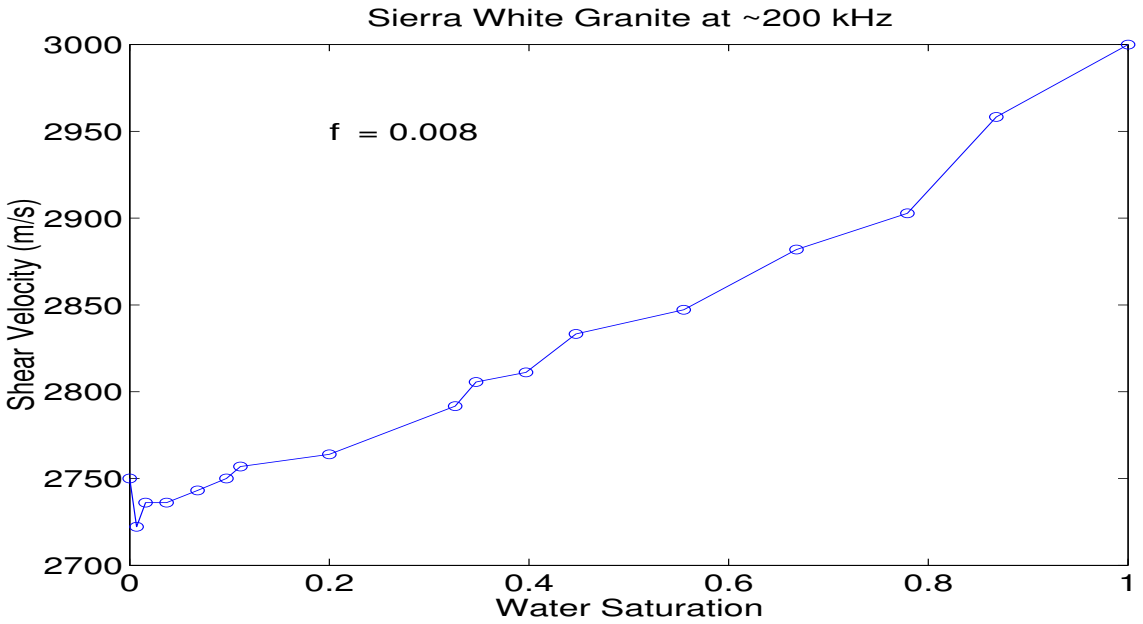


Figure 3: Shear wave velocities as a function of water saturation for drainage experiments of Murphy (1982) in Sierra White granite. jim2-swg200vs [NR]

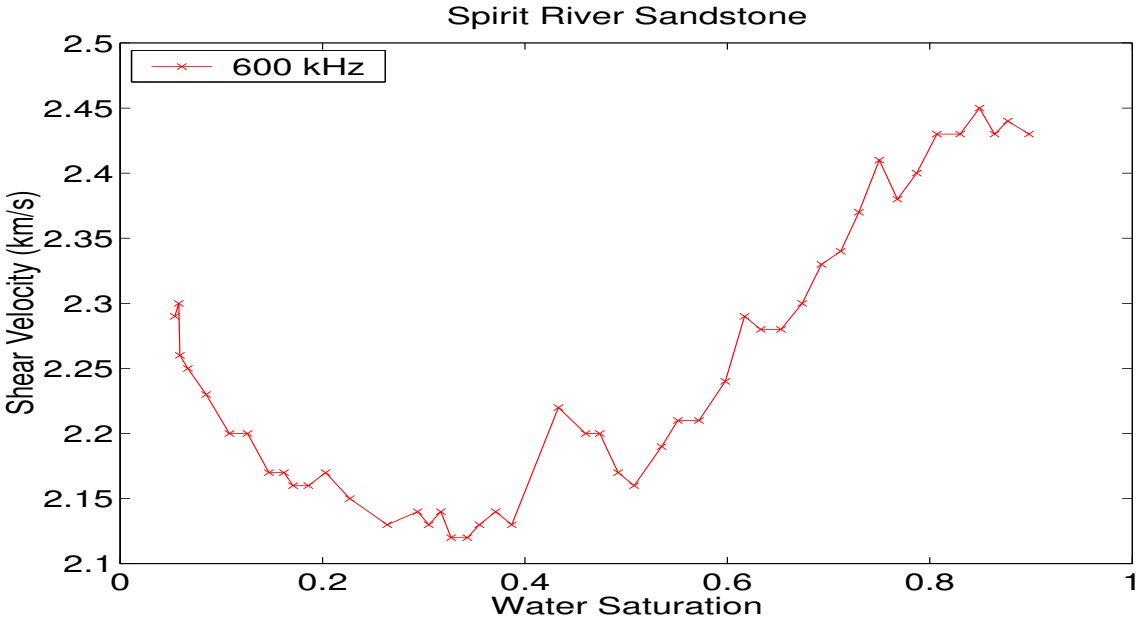


Figure 4: Shear wave velocities as a function of water saturation for drainage experiments of Knight and Nolen-Hoeksema (1990) in Spirit River sandstone. jim2-knight_vs [NR]

imaginary parts, then to a reasonable approximation it is the case that the curve of interest lies close to the real axis in the complex k_z^2-plane. If the imaginary parts exactly vanish, the curve reduces to a straight line on the real axis in this plane. These observations suggest that it might be helpful to trace rays in the complex plane radiating out from the origin, and in particular a ray (*i.e.*, a straight line) passing through the origin and also through the point corresponding to whichever point, k_s^2 or $(k_s^*)^2$, happens to lie closest to the origin should provide a good starting point for the analysis. Another alternative is to consider the straight line that connects these two points directly, even though it would generally not also be a ray through the origin (except for the special case when there is no attenuation). Both of these alternatives have been tried.

The first alternative, a ray through the origin and then passing through the closest point k_s^2 or $(k_s^*)^2$, has the very important characteristic that the values of the dispersion function become purely imaginary in the shadow of the starting point of the curve. This fact provides a great simplification because we need the dispersion function to vanish identically – both in real and imaginary parts, and this shadow region has the nice characteristic that the real part is automatically zero. So the only remaining issue is to check where the imaginary part vanishes. This procedure is much easier to implement and to understand intuitively than trying to find the complex zeroes using something like a Newton method, which could also be implemented for this problem.

The second alternative is not as rigorous as the first, but for the case of small attenuation gives very similar results and is especially easy to implement. In this case we need only consider the line connecting the two points k_s^2 and $(k_s^*)^2$ in the complex plane. It turns out that in the two cases considered here, the real part of the dispersion function is again either zero or very small, so that it makes sense to treat this approach as an approximation to the first one in that we need only seek the points where the imaginary part vanishes. This procedure is very intuitive and examples are shown in FIGS. 5 through 8.

Massillon sandstone

For Massillon, we have the Gassmann-like situation in which the shear wave speed for the drained case is smaller than that for the fully saturated case and therefore $Re(k_s^*) < Re(k_s)$. FIG. 5 shows how the imaginary parts of the dispersion function change in this case as the real part of k_z^2 varies from $Re((k_s^*)^2)$ to $Re(k_s^2)$ (*i.e.*, from air saturated to water saturated). FIG. 5 shows four of these curves ($S = 0.2$ to 0.8). FIG. 6 was generated by completing the procedure for 19 equally spaced points in saturation S. FIG. 6 shows furthermore that the curve obtained actually fits the data for Massillon better than Gassmann does (the straight line between the end points). This is a bit of a surprise as virtually everyone (including the present authors) have often considered these data to be the best known proof of the accuracy of Gassmann's equations for partial saturation problems.

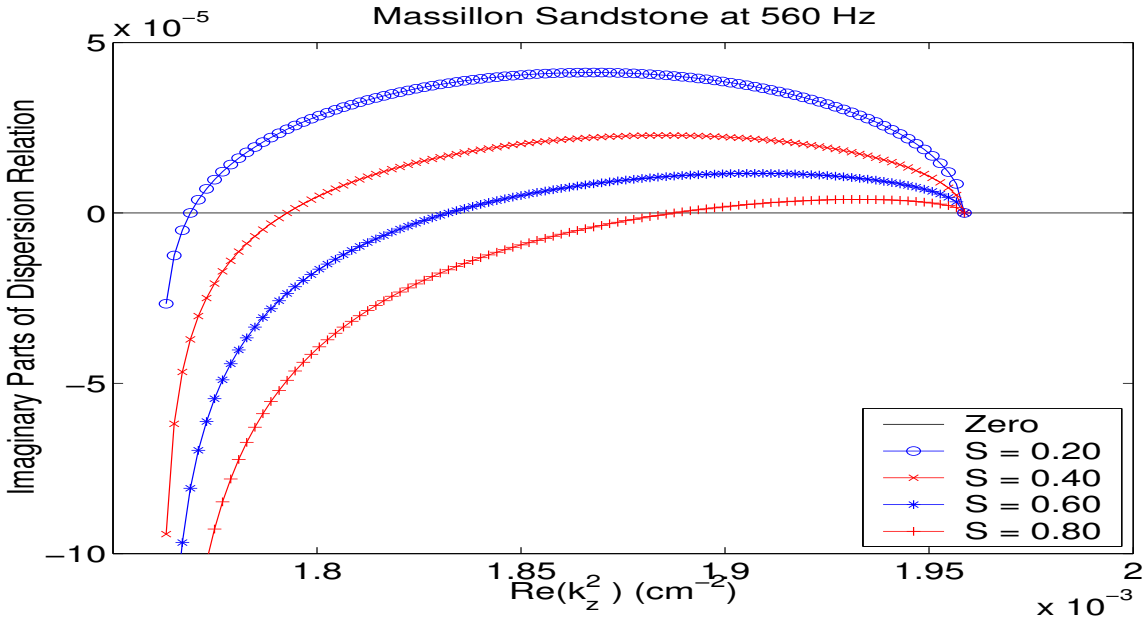


Figure 5: Showing how the imaginary parts of the dispersion relation for Massillon sandstone change in the complex k_z^2 plane as k_z varies from k_{sa} to k_{sw}. The real part of the dispersion relation is either zero or very close to zero along this line and therefore the desired points are those where the imaginary part crosses the zero line. jim2-reimdisp_mss560_20406080_ray3 [NR]

Sierra White granite

For Sierra White, we have the non-Gassmann-like situation in which the shear wave speed for the drained case is larger than that for the fully saturated case and therefore $Re(k_s^*) > Re(k_s)$. FIG. 7 shows how the imaginary parts of the dispersion function change in this case as the real part of k_z^2 varies from $Re(k_s^2)$ to $Re((k_s^*)^2)$ (*i.e.*, from water saturated to air saturated). FIG. 7 shows four of these curves ($S = 0.2$ to 0.8). FIG. 8 was generated by completing the procedure for 19 equally spaced points in saturation S. FIG. 8 shows furthermore that both data and the curve obtained here differ substantially from the simple straightline average that might have been anticipated and, furthermore, that the curve does in fact move in the right direction to agree with the data. This is also a pleasant surprise as it was certainly not known by us what to expect in this situation since our common understanding of poroelasticity does not extend to this rather difficult set of partial saturation problems.

General behavior of the curves

Since the curves obtained in FIGS. 6 and 8 are very well-behaved, it seemed like a good idea to check for simple dependencies on the saturation parameter S. Both curves look like they might be quadratic in S. This hypothesis is tested in FIGS. 9 and 10. We find that the quadratic dependence is essentially exact to graphical accuracy for Massillon sandstone, and it is close

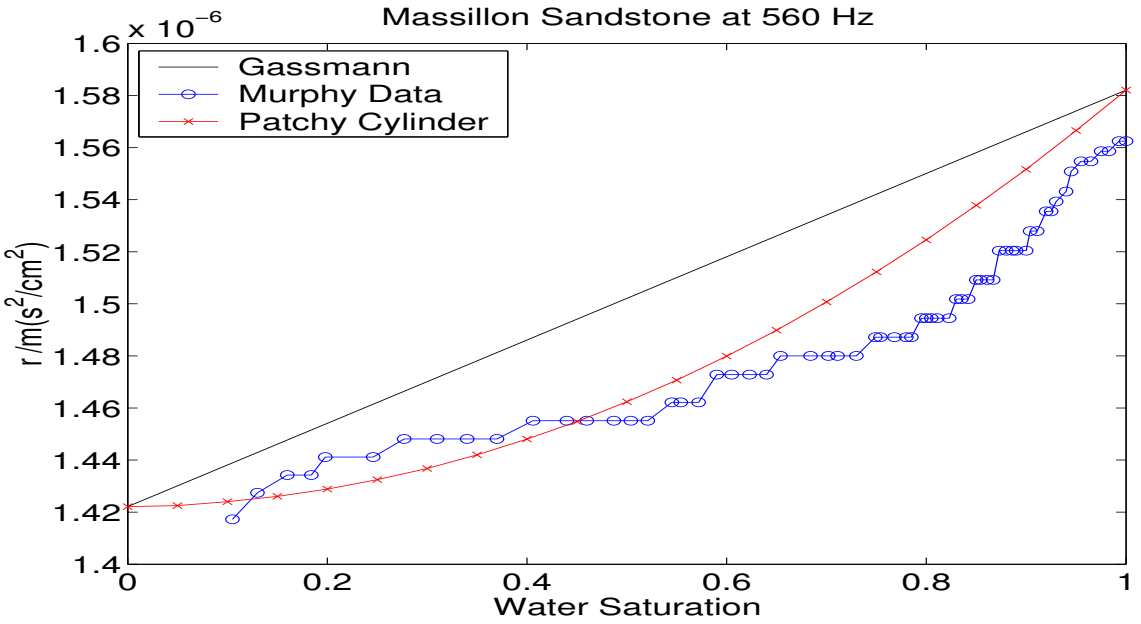


Figure 6: Comparison between the points that solve the dispersion relation for the patchy cylinder, plotted as $\rho/\mu = 1/v_s^2$ versus water saturation S, for Massillon sandstone at 560 Hz. Data are from Murphy. The Gassmann curve is computed assuming that the shear modulus μ is constant and that the only quantity chaning is therefore the density ρ. jim2-mss_rhovmu [NR]

but not exact for the Sierra White granite. We have not yet tried to analyze this behavior and will therefore not attempt an explanation of it at this time.

SUMMARY

Biot slow-wave effects in layered materials have been studied previously by Pride *et al.* (2002) and many others found in their references. The present work is motivated by the desire to understand how fluids interacting with common poroelastic systems may create viscous attenuation in partially saturated (and especially in patchy saturated) cylinders. These effects can then be observed in the attenuation of extensional and torsional waves. There are large quantities of such data already available, and one thrust of our future work will be to reanalyze these data in light of the methods developed here. We have concentrated on analysis of the wave velocities here, as this is clearly the first essential step in the overall analysis of these problems.

ACKNOWLEDGMENTS

We thank Bill Murphy and Rosemary Knight for providing access to their unpublished data files. We thank Brad Artman for helpful comments on the manuscript.

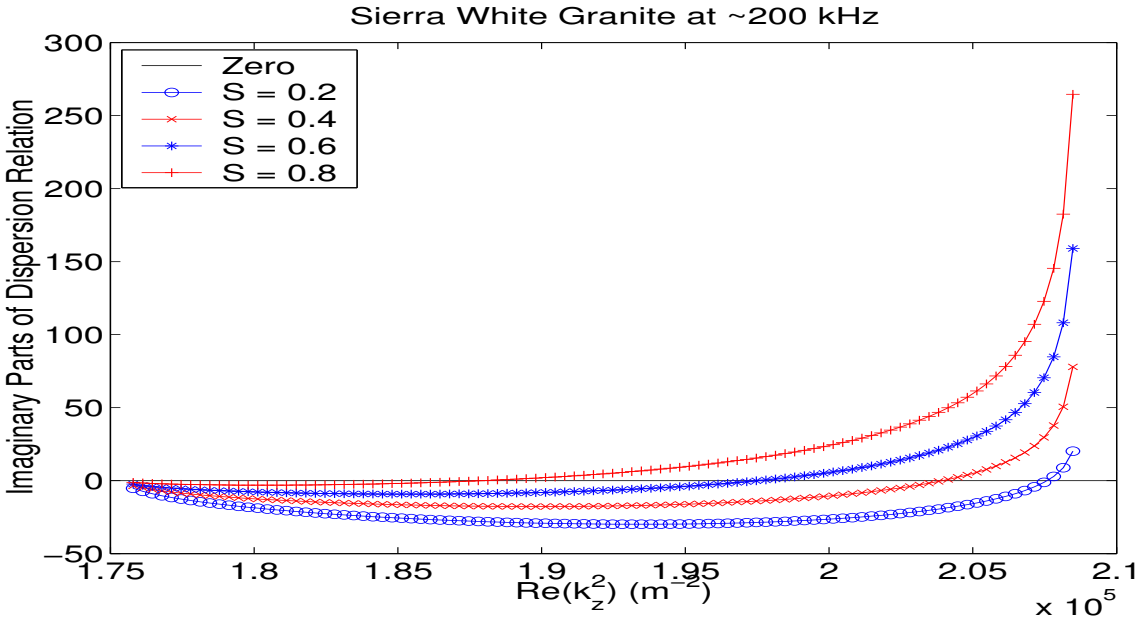


Figure 7: Showing how the imaginary parts of the dispersion relation for Sierra white granite change in the complex k_z^2 plane as k_z varies from k_{sa} to k_{sw}. The real part of the dispersion relation is either zero or very close to zero along this line and therefore the desired points are those where the imaginary part crosses the zero line. jim2-swgdisp_20406080 [NR]

REFERENCES

Abramowitz, M., and Stegun, I. A., 1965, *Handbook of Mathematical Functions*, Dover, New York, Chapter 9.

Avellaneda, M., 1987, Iterated homogenization, differential effective medium theory and applications: *Commun. Pure Appl. Math* **40**, 527–554.

Bancroft, D., 1941, The velocity of longitudinal waves in cylindrical bars: *Phys. Rev.*, **59**, 588–593.

Berge, P. A., Bonner, B. P., and Berryman, J. G., 1995, Ultrasonic velocity-porosity relationships for sandstone analogs made from fused glass beads: *Geophysics* **60**, 108–119.

Berryman, J. G., 1980a, Long-wavelength propagation in composite elastic media. I. Spherical inclusions: *J. Acoust. Soc. Am.* **68**, 1809–1819.

Berryman, J. G., 1980b, Long-wavelength propagation in composite elastic media II. Ellipsoidal inclusions: *J. Acoust. Soc. Am.* **68**, 1820–1831.

Berryman, J. G., 1983, Dispersion of extensional waves in fluid-saturated porous cylinders at ultrasonic frequencies: *J. Acoust. Soc. Am.*, **74**, 1805–1812.

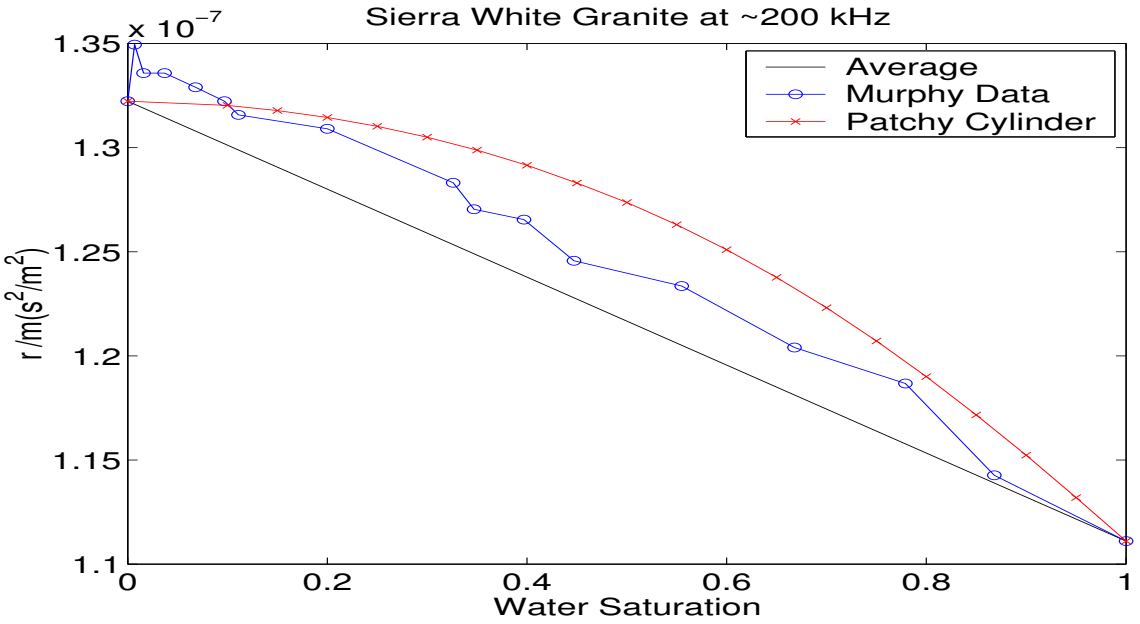


Figure 8: Comparison between the points that solve the dispersion relation for the patchy cylinder, plotted as $\rho/\mu = 1/v_s^2$ versus water saturation S, for Sierra White granite at 200 kHz. Data are from Murphy. For Sierra White, Gassmann's equation clearly does not apply since the shear modulus μ must have increased with water saturation. Data and patchy calculation results are therefore compared to the saturation weighted mean of $1/v_s^2$ in analogy to the Gassmann result. jim2-swg_rhovmu [NR]

Berryman, J. G., 1992, Single-scattering approximations for coefficients in Biot's equations of poroelasticity: *J. Acoust. Soc. Am.* **91**, 551–571.

Berryman, J. G., 1999, Origin of Gassmann's equations: *Geophysics* **64**, 1627–1629.

Berryman, J. G., Berge, P. A., and Bonner, B. P., 2000, Transformation of seismic velocity data to extract porosity and saturation values for rocks: *J. Acoust. Soc. Am.* **107**, 3018–3027.

Berryman, J. G., Berge, P. A., and Bonner, B. P., 2002a, Estimating rock porosity and fluid saturation using only seismic velocities: *Geophysics* **67**, 391–404.

Berryman, J. G., Pride, S. R., and Wang, H. F., 2002b, A differential scheme for elastic properties of rocks with dry or saturated cracks: *Geophys. J. Int.*, accepted for publication in 2002.

Berryman, J. G., Thigpen, L., and Chin, R. C. Y., 1988, Bulk elastic wave propagation in partially saturated porous solids: *J. Acoust. Soc. Am.* **84**, 360–373.

Berryman, J. G, and Wang, H. F., 2001, Dispersion in poroelastic systems: *Phys. Rev. E* **64**, 011303-1–011303-16.

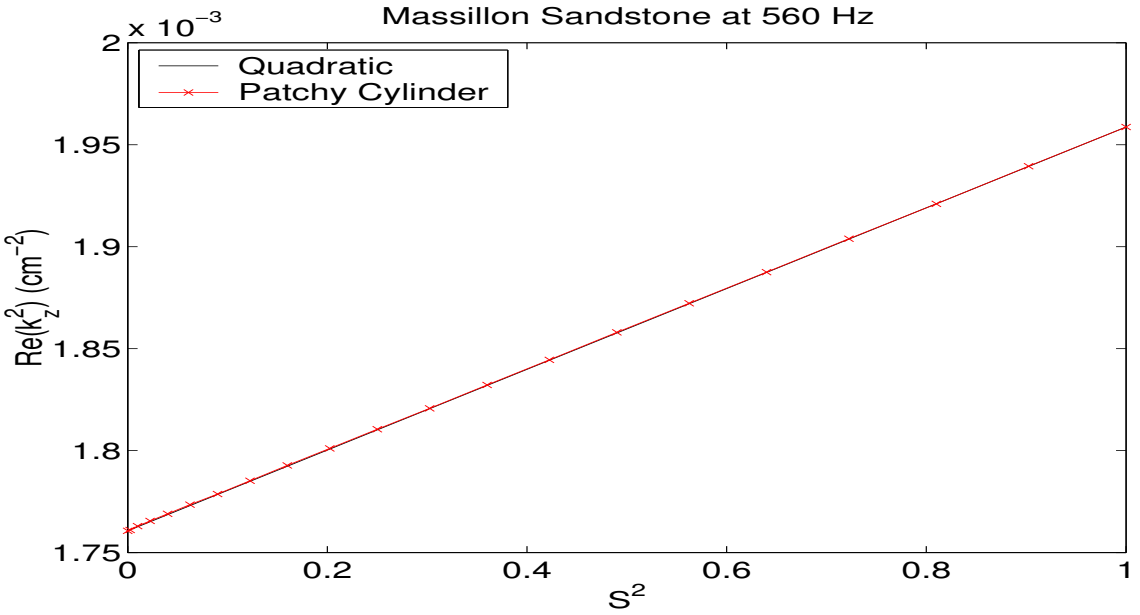


Figure 9: Comparison between the points that solve the dispersion relation in the complex k_z^2-plane, and a simple quadratic fit in water saturation S, for the Massillon sandstone measured by Murphy. The fit is nearly perfect for this case. jim2-ms_quadratic [NR]

Biot, M. A., 1956a, Theory of propagation of elastic waves in a fluid-saturated porous solid. I. Low-frequency range: *J. Acoust. Soc. Am.* **28**, 168–178.

Biot, M. A., 1956b, Theory of propagation of elastic waves in a fluid-saturated porous solid. II. Higher frequency range: *J. Acoust. Soc. Am.* **28**, 179–1791.

Biot, M. A., 1962, Mechanics of deformation and acoustic propagation in porous media: *J. Appl. Phys.* **33**, 1482–1498.

Chin, R. C. Y., Berryman, J. G., and Hedstrom, G. W., 1985, Generalized ray expansion for pulse propagation and attenuation in fluid-saturated porous media: *Wave Motion* **7**, 43–66. litem Chree, C., 1886, Longitudinal vibrations of a circular bar: *Quart. J. Pure Appl. Math.*, **21**, 287–298.

Deresiewicz, H., and Skalak, R., 1963, On uniqueness in dynamic poroelasticity: *Bull. Seismol. Soc. Am.*, **53**, 783–788.

Gassmann, F., 1951, Über die elastizität poröser medien: *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, **96**, 1–23.

Hadley, K., 1976, Comparison of calculated and observed crack densities and seismic velocities in Westerly granite: *J. Geophys. Res.* **81**, 3484–3494.

Hashin, Z., and Shtrikman, S., 1961, Note on a variational approach to the theory of composite elastic materials: *J. Franklin Inst.* **271**, 336–341.

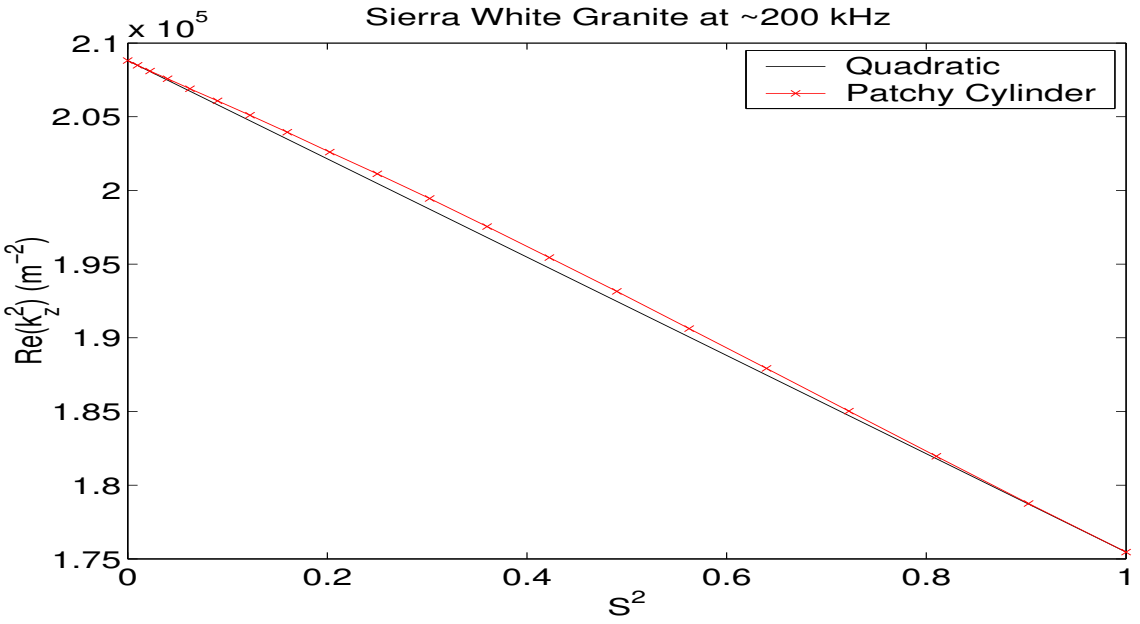


Figure 10: Comparison between the points that solve the dispersion relation in the complex k_z^2-plane, and a simple quadratic fit in water saturation S, for the Sierra White granite measured by Murphy. The fit is also good for this case, but not as good as it was for Massillon sandstone. jim2-swg_quadratic [NR]

Hashin, Z., and Shtrikman, S., 1962, A variational approach to the theory of elastic behaviour of polycrystals: *J. Mech. Phys. Solids* **10**, 343–352.

Johnson, D. L., 2001, Theory of frequency dependent acoustics in patchy-saturated porous media: *J. Acoust. Soc. Am.* **110**, 682-694.

Knight, R., and Nolen-Hoeksema, R., 1990, A laboratory study of the dependence of elastic wave velocities on pore scale fluid distribution: *Geophys. Res. Lett.* **17**, 1529–1532.

Milton, G. W., 1985, The coherent potential approximation is a realizable effective medium scheme: *Comm. Math. Phys.* **99**, 463–500.

Murphy, William F., III, 1982, *Effects of Microstructure and Pore Fluids on the Acoustic Properties of Granular Sedimentary Materials*, Ph.D. Dissertation, Stanford University.

Pochhammer, L., 1876, Über Fortpflanzungsgeschwindigkeiten kleiner Schwingungen in einem unbegrenzten isotropen Kreiscylinder: *J. Reine Angew. Math. (Crelle)*, **81**, 324–336.

Pride, S. R., and Flekkoy, E. G., 1999, Two-phase flow through porous media in the fixed-contact-line regime: *Phys. Rev. E*, **60**, 4285–4299.

Pride, S. R., Tromeur, E., and Berryman, J. G., 2002, Biot slow-wave effects in stratified rock: *Geophysics*, **67**, 271–281.

Stanford Exploration Project, Report SEP–112, September 27, 2002, pages 259–269

Exploring the relationship between uncertainty of AVA attributes and rock information

Weitian Chen and Robert Clapp[1]

ABSTRACT

Amplitude versus Angle (AVA) attributes include information about rock properties. Using a dataset from South America, we performed a multiple realization method to get multiple equal-probable AVA intercepts, gradients, and their products. We generated a 3-D histogram to evaluate the variability of those AVA attributes. In the same area, we chose a 2-D section by matching it to three wells. Then we calculated the shale volume along these three wells and found the well with low shale volume has high AVA uncertainty, which made us guess the low shale/sand ratio may cause high AVA uncertainty. The further work need to be done is to use more real data to exam our conjecture, namely, whether there exist an empirical relationship between AVA uncertainty and rock information, such as shale volume, impedance or velocity.

INTRODUCTION

Uncertainty is an inherent problem existing in velocity analysis. It is important for geophysicists to assess the variability of the velocity quantitatively. As an alternative to a common geostatistical method (Isaaks and Srivastava, 1989), Clapp (2000; 2001) introduced multiple realization method for complex operators. Clapp modified the standard geophysical inversion technique by adding random noise into the model styling goal to achieve multiple realizations. By comparing and contrasting the equal-probable realizations, the variability can be evaluated. Since the subsurface image is obtained based on the new velocity model, the uncertainty of velocity model will cause the uncertainty of amplitude information we can acquired from image (Mora and Biondi, 2000). Using the multiple realization method, Clapp (2002) showed how the velocity uncertainty affected the amplitude information.

Amplitudes carry important information about rock properties. Amplitude variation with offset (AVO) is a widely used technique in petroleum industry because AVO anomalies often indicates hydrocarbon existence. A good review of AVO analysis is provided by Castagna (1993a). Since AVO is dependent on intrinsic rock parameters such as compressional-wave velocity, shear-wave velocity, density, anisotropy and attenuation, AVO can be used to assess information for rock properties, such as lithology, porosity and pore fluid content. Castagna (1993b) provide a rock physics framework for AVO analysis.

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

The relationship between AVO and rock properties make us guess there may exist empirical relationships between AVO uncertainty and rock information. For example, if we get high variance of AVO attributes (which can be evaluated from multiple realizations) at specific subsurface areas, we can conjecture that there may be some change in rock information in the same area, such as impedance, velocity or shale/sand ratio. In this paper, we explored such relationships.

Instead of extracting amplitude variations with offset, we adopted amplitude variation with angle (AVA) analysis because realistic velocities usually break the simple relationship between offset and angle. The dataset we used was from South America. We evaluated the variability of AVA attributes by using a 3-D histogram. A 2-D section was extracted and shale volume along the wells in this section were calculated. We found the well with low shale volume has obvious higher AVA uncertainty than other two wells, which made us conjecture the low shale/sand ratio will cause high AVA uncertainty. The further work need to be done is to use more real data to exam whether our guess is true or there exist other empirical relationships between AVA uncertainty and rock information, such as shale volume, impedance or velocity.

METHODOLOGY TO EVALUATE THE UNCERTAINTY OF AVA ATTRIBUTES

We will use multiple equal-probable velocity models to get multiple images. From those images, we can extract angle gathers and get intercept A and slope B (Gratwick, 2001). By comparing and contrasting the multiple realizations of these AVA attributes, we can access their variability.

Using multiple realization method to get multiple equal-probable velocity models

Regularized geophysical inversion problems include two fitting goals: data fitting and model styling. They can be written as:

$$0 \approx \mathbf{r_d} = \mathbf{d} - \mathbf{Lm} \tag{1}$$

$$0 \approx \mathbf{r_m} = \epsilon \mathbf{Am} \tag{2}$$

An ideal regularization operator $\mathbf{A}$ should be the inverse model covariance. In practice, according to the difficulty to get the explicit model covariance, $\mathbf{A}$ is usually approximated as Lapacian, PEF or steering filter.

Generally, the regularization operator only describes the two point statistics. The first order statistics, spatial variance, is not included in it. Like in geostatistics, we can add normal noise vector $\boldsymbol{\eta}$ into model styling goal so that we can get the comparable variance in poorly determined regions as in well determined regions (Claerbout, 1999; Clapp, 2000).

The fitting goals including both first and second model statistics can be written as:

$$0 \approx \mathbf{r_d} = \mathbf{d} - \mathbf{Lm} \tag{3}$$

$$\sigma_m \boldsymbol{\eta} \approx \mathbf{r_m} = \epsilon \mathbf{Am} \tag{4}$$

Scalar σ_m can be approximated as the variance of the model residual acquired by applying regularization operator $\mathbf{A}$ to first estimated model (Claerbout, 1999). By changing normal noise $\boldsymbol{\eta}$, we can get equal-probable models from which we can evaluate the variability of the model.

When we perform velocity analysis, the data we used are the value picked from semblance. So there also exist data uncertainty. Similar to the modification of model styling goal, we can add normal noise into the data fitting goal in terms of noise covariance to include this effect on our model evaluation. The noise inverse covariance can be approximated as the chain of a diagonal operator and a PEF on $\mathbf{r}_d$. A detailed discussion on how to include data and model uncertainty to evaluate velocity was given by Clapp (2002).

Amplitude Balance

During a seismic survey, it can't be guaranteed that all the receivers have the same response. Neither can we promise that the energy for all shots have same energy. So, the recorded amplitude difference between traces will include not only rock information but also artifacts caused by different sources and receivers. We should remove such artifacts before performing AVA analysis.

For each trace, its amplitude square A_t can be expressed as

$$A_t = R_t S_t G_t U_t \tag{5}$$

Here, R_t is the component from the unique receiver the trace is corresponding to; S_t is from the unique source the trace is corresponding to; G_t corresponding to geometrical spreading for this trace; and U_t is the remaining amplitude component which contains the information related to rock properties. The way we balanced the amplitude is the following:

1. Extract the same depth window for all traces and calculate the sum of the amplitude square for all traces in this window.
2. Using the conjugate gradient method to solve for R_t, S_t, G_t and U_t using (5).
3. The problem will be underdetermined. We regularize the problem by applying Laplacian to G_t and U_t because they should be spatially continuous.
4. After solving R_t, S_t, G_t and U_t for each trace, we will divide all traces by their corresponding R_t, S_t and G_t to remove the artifacts.

Extract AVA attributes and evaluate their variability

After amplitude balancing, we perform split-step phase shift migration and extract multiple angle gathers from the image using multiple equal-probable velocity models. Then we follow Gratwick's algorithm (Gratwick, 2001) to get intercept and gradient. After extracting intercept

A and gradient B from angle gathers, we first cross-plot A and B (Ross, 2000; Castagna and Swan, 1997) and muted those clusting points in the plots which correspond to background values. Then we transformed the muted plots back to two panels A and B. A negative A*B value often is a hydrocarbon indicator. The variability of AVA attributes can be assessed by comparing and contrasting their multiple realizations. A convenient way to analysis the variability is to use histogram. Because the count used here is the function of CMPX, depth and magnitude of AVA attributes, the histogram used here is a 3-D cube.

THE SEISMIC DATA AND THE WELLS

The land seismic data used in this paper is from South America. In CMPX-CMPY coordinate, the CMPX range of seismic data is approximately from 1.65km to 5.8km and CMPY is approximately from 0.375km to 7.925km.

There are about 30 wells drilled in this area. Some of them are productive, some are moderately productive and one is dry. Most of the wells have spontoneous-potential, resistivity, gamma ray and sonic logs. Some wells have density and neutron logs.

Figure 1 shows the well locations in this area. The straight line is where we will extract seismic amplitude information. The line was chosen by matching the wells A, B and C using linear least square method, resulting in coordinate as cross-line=106.33 (CMPY=2.225km). Table 1 is the detailed information of the coordinates of three wells used to match the line.

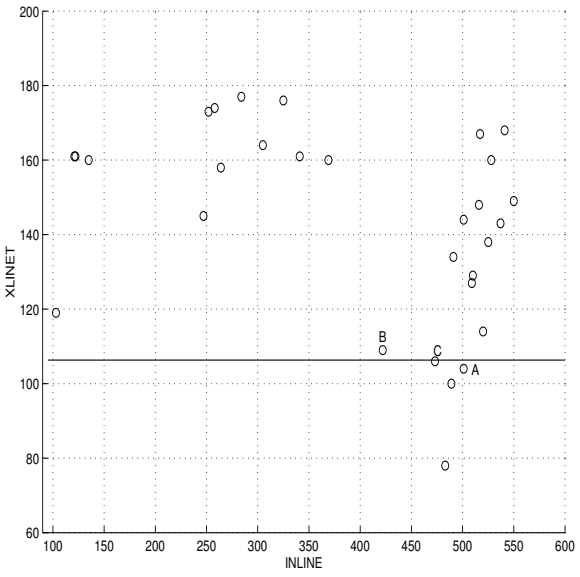


Figure 1: The well locations and the 2-D line along which we performed velocity analysis. chen-wellloc [ER]

Table 1: The coordinates of the wells

Well	Inline	Crossline	CMPX	CMPY
Well A	501	104	3.875km	2.125km
Well B	422	109	5.375km	2.375km
Well C	473	106	4.575km	2.225km

The wells don't lie exactly along the line. The cross-line distance between the line and the well A, B and C are 100m, 150m and 0 respectively. We assume the rock properties don't change rapidly across these small spans.

TRADITIONAL RESULTS

The gamma ray log is one of the most reliable logs. The well A, B and C have different depth ranges for the gamma ray log. Table 2 shows the depth range of the gamma ray log of these three wells. The depth origin is at surface, which is same as that for seismic data.

Table 2: The gamma ray log interval of the wells

Well	log start	log end
Well A	3.117 km	3.577 km
Well B (Part1)	2.033 km	2.570 km
Well B (Part2)	3.100 km	3.580 km
Well C	1.831 km	3.599 km

The gamma ray log is most frequently used to quantify shale volume along the well (Rider, 1996). According to the consolided condition of rock property in this area, we used following empirical equation (Atlas, 1992) to calculate shale volume from gamma ray value:

$$V_{shale} = 0.33\left(2^{2V^*_{shale}} - 1\right) \tag{6}$$

Here V^*_{shale} is the linear function of γ:

$$V^*_{shale} = \frac{\gamma - \gamma_{min}}{\gamma_{max} - \gamma_{min}} \tag{7}$$

The γ here is the gamma ray value, γ_{min} and γ_{max} are the minum and maximum value of the gamma ray. Figure 2 show the gamma ray log of these three wells. Figure 3 is the shale volume calculated by using equation (6).

RESULTS OF VARIABILITY STUDY

The scatterplot between AVA attributes and shale volume

We extracted AVA attributes (intercept, gradient, and their product) at three well locations: CMPX=3.875km for well A, 5.375km for well B and 4.575km for well C. The well log has much higher vertical resolution than seismic data, so, in order to correlate the AVA attributes and log data at same depth, we used sinc function to interpolate the log data. I applied the same depth window that ranges from 3.12 km to 3.57 km for the three wells and scatterploted AVA attributes and shale volume. Figure 4 show the result. We can't see obvious correlation

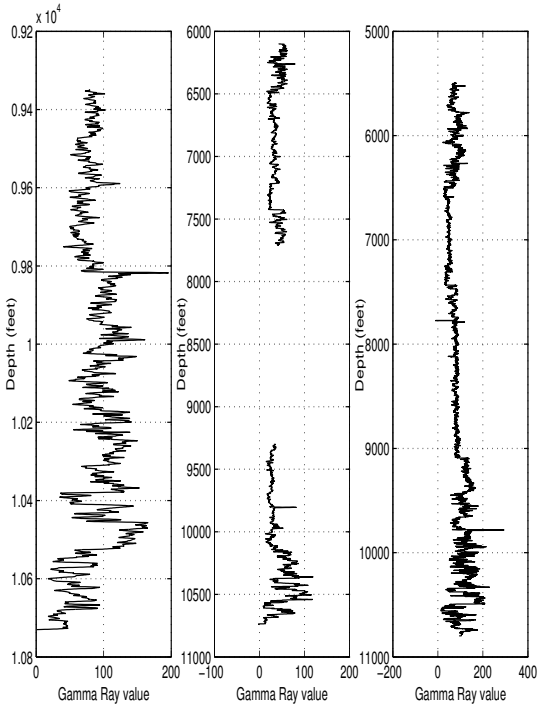


Figure 2: From left to right are the gamma ray log for the well A, B and C chen-gammaray [ER]

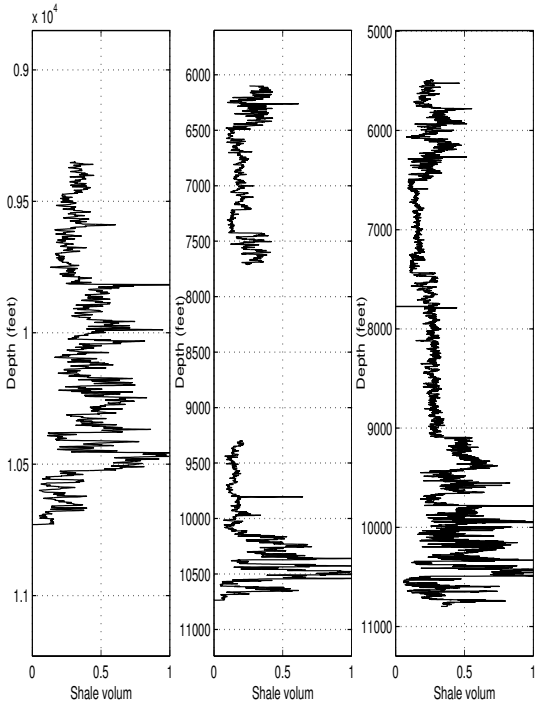


Figure 3: From left to right are the shale volume for the well A, B and C. The shale volume was calculated from gamma ray value using equation (6). I selected γ_{max} for A,B and C as 165,105 and 200 respectively by hand chen-shaleline1 [ER]

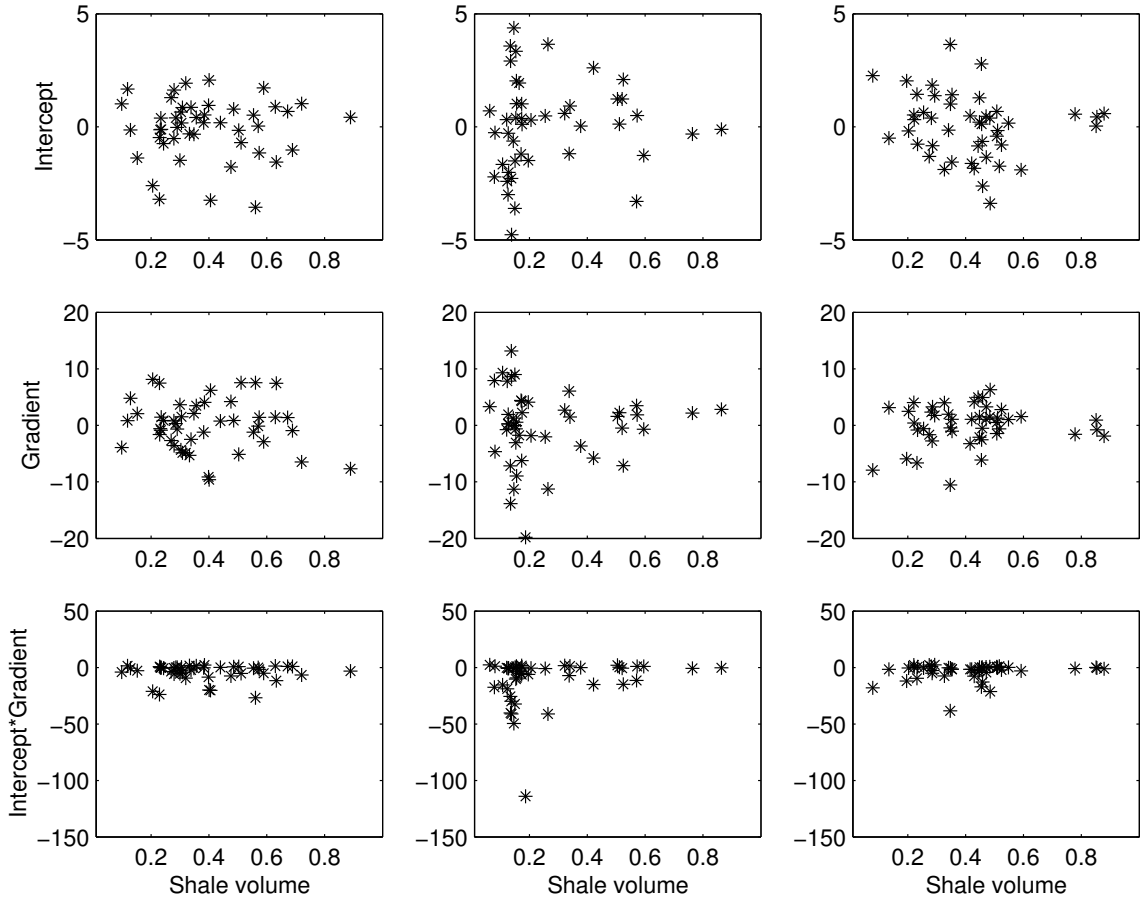


Figure 4: Scatterplot between AVA attributes and shale volume. From left to right, the column is for well A, B and C, respectively; from up to bottom, the y-axis is intercept, gradient and their product respectively. The x-axis is shale volume for all of them. chen-scatter [ER]

between AVA attributes and shale volume in this figure, but we can tell that the shale volume in the depth window along well B has a much lower value than other two wells.

Theoretically, AVA attributes will correlate better with impedance rather than shale volume. Unfortunately, we didn't have density log for well A and B. We scatterploted velocity from sonic log and AVA attributes for all three wells. We didn't see any positive caused high AVA uncertainty here.

CONCLUSION

Using a dataset from South America, we generated a 3-D histogram to conveniently evaluate the uncertainty of AVA attributes in this area. The variability of the AVA attributes was assessed using the multiple realizations method developed by Clapp (2002). From the gamma ray log we found the well with low shale volume has high AVA uncertainty, which made us guess the

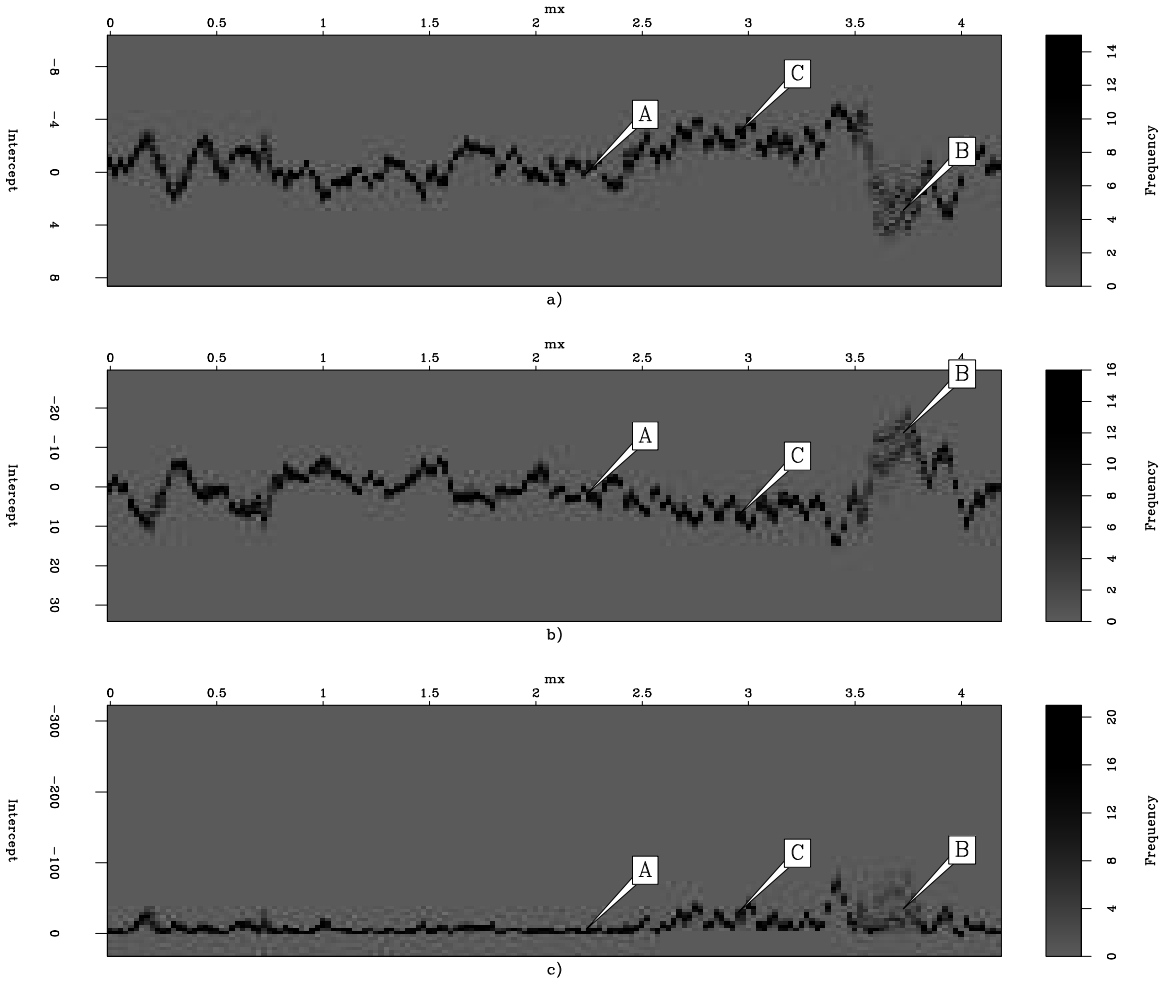


Figure 5: The depth slice of histogram cube at 3.20km. The origin of CMPX in this plot is at CMPX=1.65km chen-320.ann [ER]

low shale/sand ratio may cause high AVA uncertainty. Further work should be done using more real data to explore empirical relationships between AVA uncertainty and rock information, such as shale/sand ratio, impedance or velocity.

ACKNOWLEDGMENTS

We would like to thank the company who provided the data used in this paper; to Andres Mantilla and Tapan Mukerji for their helpful discussions; and to the SEP group for their overall support.

REFERENCES

Atlas, 1992, Digital circumferential borehole imaging log: Western Atlas International.

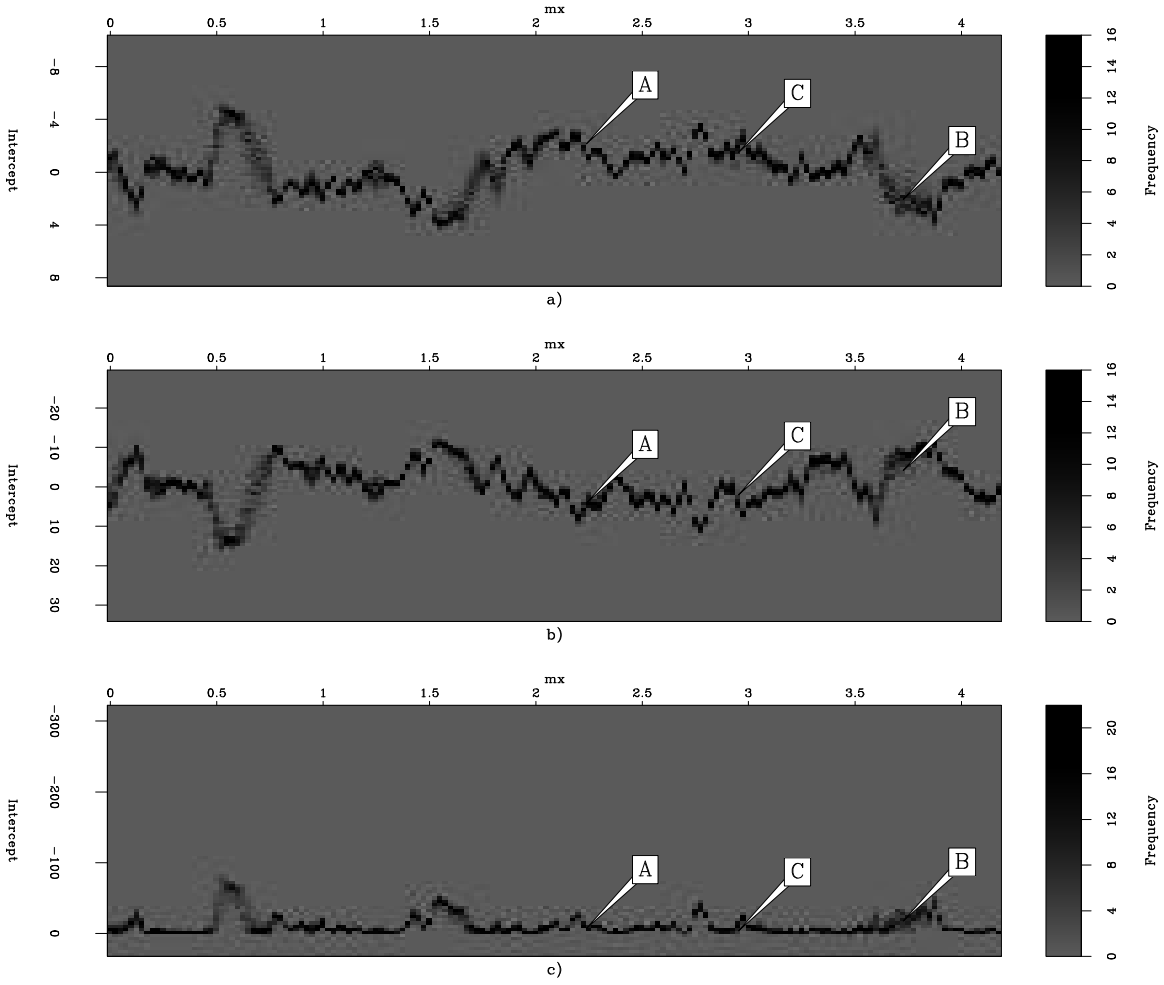


Figure 6: The depth slice of histogram cube at 3.43km. The origin of CMPX in this plot is at CMPX=1.65km chen-343.ann [ER]

Castagna, J. P., and Swan, H. W., 1997, Principles of AVO crossplotting: AAPG Mid-Continent Section Meeting; Abstracts; AAPG Bulletin, **81**, 1348.

Castagna, J. P., 1993a, Avo analysis - tutorial and review: Offset-dependent reflectivity - Theory and practice of AVO analysis.

Castagna, J. P., 1993b, Rock physics - the link between rock properties and avo response: Offset-dependent reflectivity - Theory and practice of AVO analysis.

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

Clapp, R., 2000, Multiple realizations using standard inversion techniques: SEP–**105**, 67–78.

Clapp, R. G., 2001, Multiple realizations: Model variance and data uncertainty: SEP–**108**, 147–158.

Clapp, R. G., 2002, Effect of velocity uncertainty on amplitude information: SEP–**111**, 255–269.

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

Isaaks, E. H., and Srivastava, R. M., 1989, An Introduction to Applied Geostatistics: Oxford University Press.

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

Rider, M., 1996, The geological interpretation of well logs: Gulf Publishing Company.

Ross, C. P. Effective AVO crossplot modeling: A tutorial:, 2000.

SEP ARTICLES PUBLISHED OR IN PRESS

Biondi, B., and Shan, G., 2002, Prestack imaging of overturned reflections by reverse time migration: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Biondi, B., and Vlad, I., 2002, Amplitude preserving prestack imaging of irregularly sampled 3-d data: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Biondi, B., 2001, Kirchhoff imaging beyond aliasing: Geophysics, **66**, 654–666.

Biondi, B., 2002, Stable wide-angle Fourier finite-difference downward extrapolation of 3-D wavefields: Geophysics, **67**, no. 3, 872–882.

Brown, M., 2002, Least-squares joint imaging of primaries and multiples: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Chemingui, N., and Biondi, B., 2002, Seismic data reconstruction by inversion to common offset: Geophysics, in press, **67**, no. 5, 1575–1585.

Clapp, R. G., Biondi, B., and Claerbout, J. F., 2002, Geologically constrained migration velocity analysis: submitted for publication in Geophysics.

Clapp, R. G., 2002, Effect of migration velocity uncertainty on amplitude information: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Fomel, S., Clapp, R., Prucha, M., and Berryman, J., 2001, Iterative resolution estimation in least-squares Kirchhoff migration: submitted for publication in Geophysical Prospecting.

Guitton, A., and Symes, W., 2001, Robust inversion of seismic data using the huber norm: submitted for publication in Geophysics.

Guitton, A., 2002a, Coherent noise attenuation using inverse problems and prediction error filters: First Break, **20**, 161–167.

Guitton, A., 2002b, Shot-profile migration of multiple reflections: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Prucha, M. L., and Biondi, B. L., 2002, Subsalt event regularization with steering filters: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Rosales, D., and Biondi, B., 2002, Stolt residual migration for converted waves: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

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., 2002a, Angle-domain common image gathers by wavefield continuation methods: submitted for publication in Geophysics.

Sava, P., and Fomel, S., 2002b, Wave-equation migration velocity analysis beyond the Born approximation: 72nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.

Sava, P., 2002, Prestack residual migration in the frequency domain: accepted for publication in Geophysics.

Vlad, I., 2002a, Enhanced attenuation of spatially uncorrelated noise: Enhanced attenuation of spatially uncorrelated noise:.

Vlad, I., 2002b, Reproducibility in computer-intensive sciences: Ad Astra, **1(2)**, accepted for publication.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Alvarez, Gabriel	724-0461	gabriel
Artman, Brad	723-6007	brad
Berryman, James	723-1250	berryman
Biondi, Biondo	723-1319	biondo
Brown, Morgan	723-6006	morgan
Chen, Weitian	725-1334	chen
Claerbout, Jon	723-3717	jon
Clapp, Robert	724-0461	bob
Curry, William	723-1250	bill
Guitton, Antoine	723-6007	antoine
Lau, Diane	723-1703	diane
Lomask, Jesse	724-0461	lomask
Prucha, Marie	724-0461	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	BP	(281) 366-4604	abmar1@bp.com
Francois Audebert	CGG	33 (1) 64473324	faudebert@cgg.com
Biondo Biondi	SEP	(650) 723-1319	biondo@sep.stanford.edu
Luis Canales	WesternGeco	(713) 689-5717	luiscanales@westerngeco.com
Jon Claerbout	SEP	(650) 723-3717	jon@sep.stanford.edu
Richard Cook (Co-chair, 1st year)	Shell	(713) 245-7195	cook@shellus.com
Stewart Levin	Landmark Graphics	(303) 265-8649	salevin@lgc.com
Zhiming Li (Co-chair, 2nd year)	Unocal	(281) 287-5943	zhiming.li@unocal.com
Reinaldo Michelena	Intevep	58 (212) 9086855	michelenar@pdvsa.com
Simon Spitz	CGG	(281) 646-2502	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.

Weitian Chen received a B.A. in geophysics from Peking University in 1998, and obtained a M.Sc. in geophyiscs from Peking University in 2001. He is currently working towards a Ph.D in geophysics at Stanford, and is a member of SEG.

Jon F. Claerbout (M.I.T., B.S. physics, 1960; M.S. 1963; Ph.D. geophysics, 1967), professor at Stanford University, 1967. Best Presentation Award from the Society of Exploration Geophysicists (SEG) for his paper, *Extrapolation of Wave Fields.* Honorary member and SEG Fessenden Award "in recognition of his outstanding and original pioneering work in seismic wave analysis." Founded the Stanford Exploration Project (SEP) in 1973. Elected Fellow of the American Geophysical Union. Authored three published books and five internet books. Elected to the National Academy of Engineering. Maurice Ewing Medal, SEG's highest award. Honorary Member of the European Assn. of Geoscientists & Engineers (EAGE). EAGE's highest recognition, the Erasmus Award.

Marie Clapp , formerly Marie Prucha, received her B.Sc. 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 married one of her fellow SEPers in 2001 and finally changed her last name in the summer of 2002. She is currently edging towards a Ph.D. in geophysics. She 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.

John Etgen works for BP in Houston. He spends his time taking good ideas from SEP (and a few other places), coding them up, and processing huge surveys with them. He finally learned to love the 45 degree equation just last year.

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.

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.

William W. Symes received degrees in Mathematics from the University of California at Berkeley and Harvard University, and has held faculty and research positions at the University of British Columbia, the University of Wisconsin, and Michigan State University. Since 1983 he has been Professor of Computational and Applied Mathematics at Rice University, where he also directs The Rice Inversion Project, an industry-university consortium project on mathematical and computational problems in seismic imaging.

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 A. Morton
email: morton@hess.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

Aramco Services Company
Saudi Aramco
Geophysical R&D Division
P.O. Box 750
Dhahran 31311
SAUDI ARABIA
tel: 966 (3) 874 7262
fax: 966 (3) 873 1020
contact: Mohammed N. Alfaraj
email: farajmn@aramco.com.sa

Chinese Petroleum Corporation (CPC)
Geophysical Data Processing Center
Taiwan Petroleum Exploration Division
14th Fl., No. 3, Sungren Road
Shinyi Chiu
Taipei, TAIWAN 11010
tel: 886-2-87259910
fax: 886-2-87899060
contact: Jenyang Lin
email: jenyang@cpc.com.tw

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 T. Etgen
email: etgenjt@bp.com

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

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

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) 431-6326
contact: Thomas A. 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 E. Lumley
email: david.lumley@4thwaveimaging.com

GX Technology Corporation
5847 San Felipe
Suite 3500
Houston, TX 77057-3010
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
1805 Shea Center Dr.
Suite 400
Highlands Ranch, CO 80129-2258
U.S.A.
tel: (303) 779-8080
fax: (303) 796-0807
contact: Stewart A. 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
Sector El Tambor
Edg Sede Central, FA 3306
Los Teques, Miranda 1201
VENEZUELA
Tel: 58 (212) 9087871
Fax: 58 (212) 9087633
contact: Hermes Malcotti
email: malcottih@pdvsa.com

Shell International E&P Inc.
Geophysical Technology
P. O. Box 481
Houston, TX 77001-0481
U.S.A.
tel: (713) 245-7285
fax: (713) 245-7339
contact: Chris T. Corcoran
email: ctcorcoran@shellus.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

PGS Research
10550 Richmond Ave.
Houston, TX 77042
U.S.A.
tel: (713) 735-6315
fax: (713) 532-6774
contact: James R. Myron
email: jim.myron@pgs.com

TotalFinaElf
800 Gestner
Suite 700
Houston, TX 77024
U.S.A.
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: N. Daniel Whitmore, Jr.
email: ndwhitm@ppco.com

Unocal
14141 Southwest Frwy
Sugar Land, TX 77478
U.S.A.
tel: (281) 287-7481
fax: (281) 287-5360
contact: Philip S. 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 L. Canales
email: luiscanales@westerngeco.com