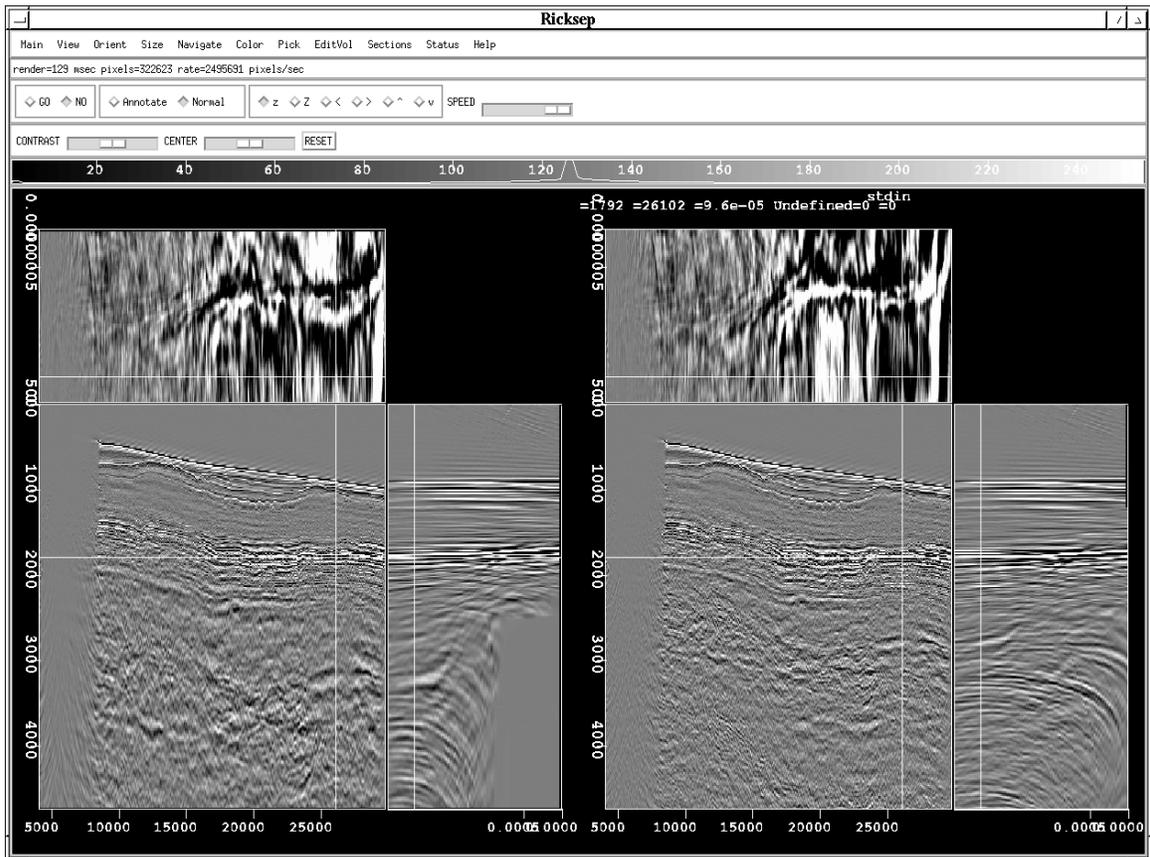


STANFORD EXPLORATION PROJECT

*Brad Artman, James Berryman, Biondo Biondi, Morgan Brown, Robert Clapp, Barbara Cox,
William Curry, Joe Dellinger, Andrey Karpushin, Clement Kostov, Marie Prucha,
Daniel Rosales, Paul Sava, and Ioan Vlad*

Report Number 110, August 2001



Copyright © 2001

by the Board of Trustees of the Leland Stanford Junior University

Stanford, California 94305

*Copying permitted for all internal purposes of the Sponsors of
Stanford Exploration Project*

Preface

The electronic version of this report¹ makes the included programs and applications available to the reader. The markings [ER], [CR], and [NR], are promises by the author about the reproducibility of each figure result.

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

ER denotes Easily Reproducible and are the results of a processing described in the paper. The author claims that you can reproduce such a figure from the programs, parameters, and makefiles included in the electronic document. The data must either be included in the electronic distribution, or be easily available (e.g. SEG-EAGE data sets) to SEP and non-SEP researchers. The data may also be available in the SEP data library, which can be viewed at http://sepwww.stanford.edu/public/docs/sepdata/lib/toc_html/. We assume you have a UNIX workstation with Fortran, Fortran90, C, X-Window system and the software downloadable from our website (SEP makerules, SEPlib, and, to properly reproduce the documents, the SEP latex package). Before the publication of the electronic document, someone other than the author tests the author's claim by destroying and rebuilding all ER figures. Some ER figures may not be reproducible by outsiders because they depend on data sets that are too large to distribute, or data that we do not have permission to redistribute but are in the SEP data library.

CR denotes Conditional Reproducibility. The author certifies that the commands are in place to reproduce the figure if certain resources are available. SEP staff have only attempted to make sure that the makefile rules exist and the source codes referenced are provided. The primary reasons for the CR designation is that the processing requires 20 minutes or more, or commercial packages such as Matlab or Mathematica.

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

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

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

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

¹<http://sepwww.stanford.edu/private/docs/sep108>

SEP-110 — TABLE OF CONTENTS

Imaging

<i>Biondi B. and Vlad I.</i> , Amplitude preserving prestack imaging of irregularly sampled 3-D data	1
<i>Brown M., Biondi B., and Kostov C.</i> , Wave-equation prestack depth migration for sub-basalt P and converted wave imaging	19
<i>Berryman J. G.</i> , Time-reversal acoustics and maximum-entropy imaging	31
<i>Rosales D., Sava P., and Biondi B.</i> , Stolt residual migration for converted waves	49
<i>Vlad I. and Biondi B.</i> , Effective AMO implementation in the log-stretch, frequency-wavenumber domain	63
<i>Rosales D. and Biondi B.</i> , On asymmetric Stolt residual migration for imaging under salt edges	71

Velocity

<i>Cox B. E.</i> , Data dependent parameterization and covariance calculation for inversion of focusing operators	75
<i>Sava P. and Biondi B.</i> , Born-compliant image perturbation for wave-equation migration velocity analysis	91
<i>Clapp R. G.</i> , Ray-based tomography with limited picking	103

Filtering

<i>Curry W. and Brown M.</i> , A new multiscale prediction-error filter for sparse data interpolation	113
<i>Brown M.</i> , Estimation of systematic errors in tracked datasets using least squares crossing point analysis	123
<i>Karpushin A. and Brown M.</i> , Whitening track residuals with PEFs in IRLS approach to the Sea of Galilee	133

AVA

<i>Artman B.</i> , 5th dimension warning for 4D studies	141
---	-----

Artman B., AVA attributes (2 for 1 special) 153

Computing

Clapp R. G., Ricksep: Interactive display of multi-dimensional data 163

Vlad I., Using the NTSC color space to double the quantity of information in an image ... 173

Manual

Clapp R., Prucha M., Sava P., Dellinger J., and Biondi B., SEP manual 181

SEP article published or in press, 2000-01 333

SEP phone directory 335

Research personnel 337

SEP sponsors for 2000-01 341

Amplitude preserving prestack imaging of irregularly sampled 3-D data

Biondo Biondi and Ioan Vlad¹

ABSTRACT

We introduce a computationally efficient and robust method to regularize acquisition geometries of 3-D prestack seismic data before prestack migration. The proposed method is based on a formulation of the geometry regularization problem as a regularized least-squares problem. The model space of this least-squares problem is composed of uniformly sampled common offset-azimuth cubes. The regularization term fills the acquisition gaps by minimizing inconsistencies between cubes with similar offset and azimuth. To preserve the resolution of dipping events in the final image, the regularization term includes a transformation by Azimuth Moveout (AMO) of the common offset-azimuth cubes. The method is computationally efficient because we applied the AMO operator in the Fourier-domain, and we precondition the least-squares problem. Therefore, no iterative solution is needed and excellent results are obtained by applying the adjoint operator followed by a diagonal weighting in the model domain.

We tested the method on a 3-D land data set from South America. Subtle reflectivity features are better preserved after migration when the proposed method is employed as compared to more standard geometry regularization methods. Furthermore, a dipping event at the reservoir depth (more than 3 km) is better imaged using the AMO regularization as compared to a regularization operator that simply smoothes the data over offsets.

INTRODUCTION

Irregular acquisition geometries are a serious impediment to the accurate imaging of the subsurface. When the data are irregularly sampled, images are often affected by amplitude artifacts and phase distortions, even if the imaging algorithm employed is designed to preserve amplitudes. Addressing this problem becomes more crucial when the goal is to use information contained in the image amplitudes. In these cases, the application of a simple imaging sequence that relies on standard ‘adjoint’ imaging operators is likely to produce misleading results. Amplitude-preserving imaging of irregular geometries is thus one area of seismic processing that can greatly benefit from the application of inverse theory, and extensive research have been carried out in this direction.

There are two distinct approaches that can be used to apply inverse theory to the prob-

¹email: biondo@sep.stanford.edu

lem. The first one attempts to regularize the data geometry before migration (Duijndam et al., 2000), while the second one attempts to correct for the irregular geometries during migration (Albertin et al., 1999; Bloor et al., 1999; Audebert, 2000; Rousseau et al., 2000; Duquet et al., 1998; Nemeth et al., 1999). The main strength of the latter approach is also its main weakness; it is model based; in particular, it depends on an accurate knowledge of the interval velocity model. If the model is well known, the methods based on the inversion of imaging operators have the potential of being accurate, because they exploit the intrinsic correlation between seismic traces recorded at different locations. However, when the uncertainties on the model are large, these methods can be also unreliable. Furthermore, a full prestack migration is an expensive process. Its substitution with an inversion process, even if iterative or approximate, might be beyond the practical reach.

In this paper we propose a method that has the advantages of both approaches. We regularize the data geometry before migration, but to fill the acquisition gaps we use a partial migration operator that exploits the intrinsic correlation between prestack seismic traces. The imaging operator is Azimuth Moveout (AMO) (Biondi et al., 1998), that depends on a primary knowledge of RMS velocity. RMS velocity can be estimated from the data much more robustly than interval velocity.

Ronen (1987) was the first to use a partial migration operator to improve the estimate of a regularized data set. His method uses dip moveout (DMO) to regularize stacked cubes. Chemingui and Biondi (1997; 1999) have previously inverted AMO to create regularly sampled common offset-azimuth cubes. The main advantages of the method proposed in this paper over the previous methods are: a) it is based on a Fourier-domain implementation of AMO (Vlad and Biondi, 2001), as opposed to a Kirchhoff implementation, and thus it is computationally efficient and its implementation is straightforward, b) it uses AMO in the regularization equation (model styling) formulation of a regularized least-squares inverse problem, instead that in the modeling equation. In this formulation, the regularization term can be effectively preconditioned, with a substantial gain in computational efficiency, c) it approximates the solution of the preconditioned least-squares problem by applying normalization weights to the model vector after the application of the adjoint operator. Therefore, it avoids the costs and pitfalls of iterative solutions.

Our formulation of the geometry regularization problem as a regularized least-squares problem is similar to the formulation that Fomel presented in his Ph.D. thesis (2001). He uses a finite difference implementation of offset continuation where we use a Fourier implementation of AMO. These two operators are kinematically equivalent, and their computational efficiency is similar. However, the methods are different with respect to items b) and c) listed above. Our method should be more efficient because it explicitly preconditions the regularization term by inverting it. The inversion is fast because exploits the fact that the regularization matrix can be factored into the product of a block lower-diagonal matrix with a block upper-diagonal matrix, which are easily invertible by recursion. The preconditioning substantially improves the conditioning of the problem; therefore, a simple diagonal normalization of the model vector yields a good and fast solution to the problem.

NORMALIZED PARTIAL STACKING AND INVERSE THEORY

Our goal is to create uniformly sampled common offset/azimuth cubes, that can be migrated using an amplitude-preserving algorithm. The main tool to create these common offset/azimuth cubes is partial stacking the data recorded with irregular geometries within offset and azimuth ranges.

Stacking is the operation of averaging seismic traces by summation. It is an effective way to reduce the size of data sets and to enhance reflections while attenuating noise. To avoid attenuating the signal together with the noise, the reflections need to be coherent among the traces that are being stacked. A common method to increase trace coherency is to apply Normal Moveout (NMO). NMO is a first-order correction for the differences in timings among the reflections in traces recorded at different offsets. Global stacking of all the traces recorded at the same midpoint location, no matter their offset and azimuth, is the most common type of stacking. Partial stacking averages only those traces with offset and azimuth within a given range.

The first problem that we encounter when stacking 3-D prestack data is that, because of acquisition geometry irregularities, data traces do not share the same exact midpoint location. Stacking 3-D prestack data is thus the combination of two processes: spatial interpolation followed by averaging. To start our analysis we 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 it is equal to the weighted sum of the neighboring stacked traces. The interpolation weights are functions of the distance between the midpoint location of the model trace and the midpoint location of the data trace. The sum of all the weights corresponding to one data trace is usually equal to one. Because the weights are independent from time along the seismic traces, for sake of notation simplicity, we collapse the time axis and consider each element d_i of the data space (recorded data) \mathbf{d} , and each element m_j of the model space \mathbf{m} (stacked volume), as representing a whole trace. The relationship between data and model is linear and can be expressed as,

$$d_i = \sum_j a_{ij} m_j; \text{ subject to the constraint } \sum_j a_{ij} = 1. \quad (1)$$

In matrix notation, equation (1) becomes

$$\mathbf{d} = \mathbf{A}\mathbf{m}. \quad (2)$$

where the model vector \mathbf{m} is the composite of all the offset/azimuth cubes \mathbf{m}_{h_i} , that is

$$\mathbf{m} = \begin{bmatrix} \mathbf{m}_{h_1} \\ \vdots \\ \mathbf{m}_{h_i} \\ \vdots \\ \mathbf{m}_{h_n} \end{bmatrix}. \quad (3)$$

Stacking is the summing of the data traces into the model traces weighted by the interpolation weights. In operator notation, stacking can be represented as the application of the adjoint

operator \mathbf{A}' to the data traces (Claerbout, 1998); that is,

$$\mathbf{m} = \mathbf{A}' \mathbf{d}. \quad (4)$$

The application of simple stacking as an adjoint operator does not yield satisfactory results when the fold is unevenly distributed among midpoint bins. In the stack, the amplitudes of the bins with higher fold will be artificially higher than the amplitudes of the bins with lower fold. To compensate for this unevenness in the fold, it is common practice to divide the stacked traces by the inverse of the fold. This fold normalization can be also expressed in operator notation when a diagonal operator \mathbf{W}_m is added in front of the adjoint operator in the computation of the stack:

$$\mathbf{m} = \mathbf{W}_m \mathbf{A}' \mathbf{d} = \mathbf{d}. \quad (5)$$

The weights w_j^m are given by the inverse of the fold, that can be simply computed by a summation of the elements in each column of \mathbf{A} ; that is,

$$w_j^m = (\sum_i a_{ij})^{-1}. \quad (6)$$

Data gaps in the fold present a problem for fold normalization because they make the weights diverge to infinity in equation (6). To avoid instability, it is common practice to add a small number ϵ_w to the actual fold, or to set the weights to zero when the fold is smaller than ϵ_w , as in the following revised expression for the weights:

$$w_j^m = \begin{cases} (\sum_i a_{ij})^{-1} & \text{if } \sum_i a_{ij} \geq \epsilon_w \\ 0 & \text{elsewhere.} \end{cases} \quad (7)$$

We derived the fold normalization by a simple heuristic, and it may seem an *ad hoc* solution to the problem of normalizing the output of stacking. However, it can be shown that the weights used by fold normalization can be derived from applying the general theory of inverse least-squares to the stacking normalization problem (Biondi, 1999). The least-squares problem is

$$0 \approx \mathbf{d} - \mathbf{A}\mathbf{m}, \quad (8)$$

and its formal solution can be written as

$$\mathbf{m} = \mathbf{A}_m^\ddagger \mathbf{d} = \left(\mathbf{A}' \mathbf{A} \right)^{-1} \mathbf{A}' \mathbf{d}. \quad (9)$$

where the operator \mathbf{A}_m^\ddagger is often referred as the pseudoinverse (Strang, 1986). Applying the least-squares inverse is equivalent to applying the adjoint operator \mathbf{A}' followed by a *spatial filtering* of the model space given by the inverse of $\mathbf{A}' \mathbf{A}$. The fold normalization can be seen as a particular approximation of the inverse of $\mathbf{A}' \mathbf{A}$ with a diagonal operator. Because of the size of the problem, computing the exact inverse of $\mathbf{A}' \mathbf{A}$ is not straightforward. We have thus two choices: 1) to compute an analytical approximation to the inverse; 2) to use an iterative method to compute a numerical approximation to the inverse. Even if we follow the second strategy, the availability of an analytical approximation to the inverse is useful, because the approximate inverse can be used as a preconditioner to accelerate the convergence of the iterative inversion.

We will discuss two methods for approximating the inverse of $\mathbf{A}'\mathbf{A}$. The first method is algebraic and it is based on the direct manipulation of the elements of $\mathbf{A}'\mathbf{A}$, such as extracting its diagonal or summing the elements in its columns (or rows). The second method is based on the idea that for capturing the most significant properties of $\mathbf{A}'\mathbf{A}$ by measuring its effects when applied to a reference model (\mathbf{m}_{ref}) (Claerbout and Nichols, 1994; Rickett, 2001).

Although these two methods seems unrelated, they yield equivalent results for specific choices of \mathbf{m}_{ref} . Therefore, the second method can be used to analyze the assumptions that underly the possible choices of approximations.

Approximation of $\mathbf{A}'\mathbf{A}$ by the sum of its columns

We now analyze the properties of the approximate inverse that is defined by the substitution of $\mathbf{A}'\mathbf{A}$ with a diagonal matrix $\widetilde{\mathbf{A}'\mathbf{A}}$, when the diagonal elements of $\widetilde{\mathbf{A}'\mathbf{A}}$ are equal to the sum of the corresponding rows of $\mathbf{A}'\mathbf{A}$. Notice that $\mathbf{A}'\mathbf{A}$ is symmetric, and thus summing over the columns would be equivalent to summing over the rows. It can be noticed that the diagonal elements of $\widetilde{\mathbf{A}'\mathbf{A}}$ are equal to the sum of the columns of the original operator \mathbf{A} as defined in equation (1); i.e., they are equal to the fold computed for each corresponding model trace. The weights a_{lj} are interpolator weights, and they fulfill the constraint $\sum_j a_{lj} = 1$ expressed in equation (1). It immediately follows that

$$\sum_i w_i^m = (\sum_l a_{li})^{-1}, \quad (10)$$

which is equivalent to equation (6). In summary, we just derived, for the model-space inverse, an approximation that is easy to compute, and it is equal to the fold normalization that was defined by heuristic considerations. However, the definition of the approximate inverse can be applied when more complex imaging operators are involved, and for which we do not have a simple heuristic to define the weights that improve the results obtained by application of the adjoint operator.

Approximation of $\mathbf{A}'\mathbf{A}$ by application to a reference model (\mathbf{m}_{ref})

The second method is based on the idea that for capturing the most significant properties of $\mathbf{A}'\mathbf{A}$ by measuring its effects when applied to a reference model (\mathbf{m}_{ref}).

The approximation is then evaluated as

$$\mathbf{A}'\mathbf{A} \approx \frac{\text{diag}(\mathbf{A}'\mathbf{A} \mathbf{m}_{\text{ref}})}{\text{diag}(\mathbf{m}_{\text{ref}})}. \quad (11)$$

This method has the important advantage that it does not require the explicit evaluation (and storage) of the elements of $\mathbf{A}'\mathbf{A}$, but it just requires its application to a reference model. Of course, the resulting approximation is strongly dependent from the choice of the reference model. The closer is the reference model to the true model, the better is the approximation. In theory, if the reference model is equal to the true model we will achieve perfect results.

It is also easy to show that the approximation of $\mathbf{A}'\mathbf{A}$ by the sum of its columns, equation (10), is equivalent to the choice of a constant vector as \mathbf{m}_{ref} in equation (11). Therefore, it will bias the imaging process towards model that are constant. In the case of stacking, it encourages flat reflectors, that is consistent with the flat reflector assumptions underlying the stacking process. In the case of a more complex imaging operator aimed at imaging complex structure, this bias towards flat reflectors may be less appropriate.

Fold normalization is effective when the geometry is irregular but without sizable data gaps. However, when these gaps are present the normalization weights tend to become large. Even if instability can be easily avoided by the weights modification expressed in equation (7), gaps are going to be left in the uniformly sampled data. These gaps are likely to introduce artifacts in the image because migration operators spread them as *migration smiles*. The gaps should be filled using the information from nearby traces before migration. In the next section we discuss how that can be done within the context of inverse theory.

MODEL REGULARIZATION AND PRECONDITIONING

In the previous section we have seen that data gaps are a challenge for simple fold normalization. To fill the gaps, we want to use the information from traces recorded with geometry similar to the missing ones. The challenge is to devise a method that maximizes the image resolution and minimizes artifacts.

Given no *a priori* knowledge on the reflectors geometry, using the information from traces from the surrounding midpoints and same offset-azimuth range can cause a resolution loss because it may smooth true reflectivity changes. On the other hand, because of physical constraints on the reflection mechanism, the reflection amplitudes can be assumed to be a smooth function of the reflection angle and azimuth. This observation leads to the idea that smoothing the data over offset and azimuth could be performed without losing resolution. Ideally, such smoothing should be done over aperture angles (dip and azimuth) at the reflection location, not over offset and azimuth at the surface. However, smoothing at the reflectors would require a full migration of the data. The migration step would make the method dependent on the accurate knowledge of the interval velocity model. This reliance on the velocity model is inescapable when the imaging problems are caused by the complexities in velocity model itself, (e.g. subsalt illumination (Prucha et al., 2001)), but it ought to be avoided when the imaging problems are caused by irregularities in the acquisition geometries.

In the context of least-squares inversion, smoothing along offset/azimuth in the model space (e.g. uniformly sampled offset/azimuth cubes) can be accomplished by introducing a model regularization term that penalizes variations of the seismic traces between the cubes. The simple least-squares problem of equation (8) then becomes

$$\begin{aligned} 0 &\approx \mathbf{d} - \mathbf{A}\mathbf{m} \\ 0 &\approx \epsilon_D \mathbf{D}'_h \mathbf{D}_h \mathbf{m}, \end{aligned} \tag{12}$$

where the *roughener* operator \mathbf{D}_h is

$$\mathbf{D}_h = \frac{1}{1 - \rho_D} \begin{bmatrix} 1 - \rho_D \mathbf{I} & 0 & 0 & \vdots & 0 & 0 \\ -\rho_D \mathbf{I} & \mathbf{I} & 0 & \vdots & 0 & 0 \\ 0 & -\rho_D \mathbf{I} & \mathbf{I} & \vdots & 0 & 0 \\ \dots & \dots & \ddots & \vdots & 0 & 0 \\ 0 & 0 & \dots & \ddots & -\rho_D \mathbf{I} & \mathbf{I} \end{bmatrix}. \quad (13)$$

The coefficient ρ_D must be between 0 and 1. It determines the range over which we smooth the offset/azimuth cubes. Smaller the value we set for ρ_D , narrower the smoothing range is.

Regularization with a roughener operator such as $\mathbf{D}'_h \mathbf{D}_h$ has the computational drawback that it substantially worsens the conditioning of the problem, making the solution more expensive. However, the problem is easy to precondition because $\mathbf{D}'_h \mathbf{D}_h$ is easy to invert, since it is already factored in a lower block-diagonal operator \mathbf{D}_h and in an upper block-diagonal operator \mathbf{D}'_h , that can be inverted by recursion. Therefore, we can write the preconditioned least-squares problem

$$\begin{aligned} 0 &\approx \mathbf{d} - \mathbf{A} \left(\mathbf{D}'_h \mathbf{D}_h \right)^{-1} \mathbf{p} \\ 0 &\approx \epsilon_D \mathbf{I} \mathbf{p}, \end{aligned} \quad (14)$$

where $\mathbf{p} = \mathbf{D}'_h \mathbf{D}_h \mathbf{m}$ is the preconditioned model vector.

To take into account fold variations we can introduce a diagonal scaling factor, by applying the same theory discussed in the previous section. The weights for the regularized and preconditioned problem are thus computed as

$$\mathbf{W}_I^{-1} = \frac{\text{diag} \left\{ \left[\left(\mathbf{D}_h \mathbf{D}'_h \right)^{-1} \mathbf{A}' \mathbf{A} \left(\mathbf{D}'_h \mathbf{D}_h \right)^{-1} + \epsilon_D \mathbf{I} \right] \mathbf{p}_{\text{ref}} \right\}}{\text{diag}(\mathbf{p}_{\text{ref}})}. \quad (15)$$

Notice that because of the nature of $\mathbf{D}'_h \mathbf{D}_h$, $\mathbf{p}_{\text{ref}} = \mathbf{D}'_h \mathbf{D}_h \mathbf{m}_{\text{ref}} = \mathbf{D}'_h \mathbf{D}_h \mathbf{1} = \mathbf{1}$, and $\left(\mathbf{D}'_h \mathbf{D}_h \right)^{-1} \mathbf{1} = \mathbf{1}$, and some computation can be saved by computing the weights as

$$\mathbf{W}_I^{-1} = \frac{\text{diag} \left\{ \left[\left(\mathbf{D}_h \mathbf{D}'_h \right)^{-1} \mathbf{A}' \mathbf{A} + \epsilon_D \mathbf{I} \right] \mathbf{1} \right\}}{\text{diag}(\mathbf{1})}. \quad (16)$$

In this case the computational cost of applying twice the leaky integrator $\left(\mathbf{D}'_h \mathbf{D}_h \right)^{-1}$ to the model vector is small, thus the computational saving is trivial. However, when we introduce more expensive operators to smooth the data over offset/azimuth, substituting equation (15) with equation (16) halves the computational cost of evaluating the weights. The solution of the problem obtained by normalizing the preconditioned adjoint is

$$\tilde{\mathbf{m}} = \left(\mathbf{D}'_h \mathbf{D}_h \right)^{-1} \mathbf{W}_I \left(\mathbf{D}_h \mathbf{D}'_h \right)^{-1} \mathbf{A}' \mathbf{d}. \quad (17)$$

The normalization weights could be used for another preconditioning step, by using them in another change of variables $\mathbf{q} = \mathbf{W}_I^{-\frac{1}{2}} \mathbf{p}$. This would yield the following least-squares problem,

$$\begin{aligned} 0 &\approx \mathbf{d} - \mathbf{A} \left(\mathbf{D}'_h \mathbf{D}_h \right)^{-1} \mathbf{W}_I^{\frac{1}{2}} \mathbf{q} \\ 0 &\approx \epsilon_D \mathbf{I} \mathbf{W}_I^{\frac{1}{2}} \mathbf{q}. \end{aligned} \quad (18)$$

If the problem expressed in equations (18) were to be solved iteratively, it is likely that it would converge faster than either the original regularized problem [equations (12)] or the preconditioned problem [equations (14)]. However, we prefer a non-iterative solution both because it is cheaper, and because it is somewhat more predictable.

Regularization by AMO

The main drawback of the method described above is that smoothing over offset/azimuth cubes by the inverse of the simple roughener operator expressed in equation (13) may result in loss of resolution when geological dips are present. It is well known that dipping events are not flattened by NMO with the same velocity as flat events. However, the method is easily generalized by substitution of the identity matrix in the lower diagonal of \mathbf{D}_h with an appropriate operator that correctly transforms a common offset-azimuth cube into an equivalent cube with a different offset and azimuth. This can be accomplished by AMO (Biondi et al., 1998). Since the cubes to be transformed are uniformly sampled we can use a Fourier-domain formulation of AMO that is both efficient and straightforward to implement (Vlad and Biondi, 2001). The roughener operator that includes AMO is then expressed as

$$\tilde{\mathbf{D}}_h = \frac{1}{1 - \rho_D} \begin{bmatrix} 1 - \rho_D \mathbf{I} & 0 & 0 & \vdots & 0 & 0 \\ -\rho_D \mathbf{T}_{h_{1,2}} & \mathbf{I} & 0 & \vdots & 0 & 0 \\ 0 & -\rho_D \mathbf{T}_{h_{2,3}} & \mathbf{I} & \vdots & 0 & 0 \\ \dots & \dots & \ddots & \vdots & 0 & 0 \\ 0 & 0 & \dots & \ddots & -\rho_D \mathbf{T}_{h_{n-1,n}} & \mathbf{I} \end{bmatrix}, \quad (19)$$

where $\mathbf{T}_{h_{i,i+1}}$ is the AMO operator that transforms the offset-azimuth cube i into the offset-azimuth cube $i + 1$. The $\tilde{\mathbf{D}}'_h \tilde{\mathbf{D}}_h$ operator can also be easily inverted by recursion and thus the least-squares problem obtained by substituting $\tilde{\mathbf{D}}_h$ for \mathbf{D}_h in equations (12) can also be easily preconditioned and normalized using the same techniques described in equations (14-18).

IMAGING OF A 3-D LAND DATA SET

We tested the geometry regularization methods presented in the previous section on a land data set recorded in South America. The data were shot with a cross-swath geometry. The shot lines are not perfectly aligned along the perpendicular to the receiver lines, but they are

oriented at an angle (about 18 degrees). Figure 1 shows the plot of the absolute offset vs. azimuth for a random sample of traces in the data set. For land data, the data set has fairly narrow-azimuth acquisition geometry.

Figure 2 shows three sections (depth slice on the top, cross-line section in the middle, and in-line section on the right) of the near-angle stack of the migrated data. At a depth of about 3.2 km there is a producing reservoir with numerous wells that could be used in future work to evaluate the imaging results. The only a structural feature in the data is a steep folding/faulting at the reservoir level. There are subtle stratigraphic features at a shallower depth, such as the channel visible in the upper-left corner of the depth slice in the figure (depth of 1.9 km).

The processing sequence comprised of the following steps: a) NMO, b) geometry regularization, c) inverse NMO, d) 3-D prestack common-azimuth wave-equation migration, with the imaging step designed to preserve relative amplitudes, as discussed by Sava and Biondi (2001). The velocity model was fairly simple. The NMO velocity was slightly varying in the lateral direction. However, we migrated the data using an interval velocity only as a function of depth.

To test the relative performances of the different geometry regularization methods we migrated the data after regularizing the geometries with three different methods: a) normalization by partial stack fold [equation (5)]; we will simply call this method *normalization*. b) normalization of the regularized and preconditioned solution without AMO [equation (17)]; we will simply call this method *regularization*, c) normalization of the regularized and preconditioned solution with AMO [equation (17) with $\mathbf{D}_h = \tilde{\mathbf{D}}_h$, as in equation (19)]; we will call this method *AMO regularization*.

The geometry regularizations produced common offset/azimuth cubes at zero azimuth and with offset sampling of 195 m, in-line midpoint sampling of 25 m, and cross-line midpoint sampling of 50 m. By assuming reciprocity, the offset sign was ignored to increase the effective fold of the common offset/azimuth cubes. Our first tests to regularize the data geometry produced common offset-azimuth cubes at zero azimuth (i.e. the data azimuth was ignored) because of the fairly limited azimuthal range at far offsets. However, we have indications that taking into account the data azimuth for the far-offset traces may be beneficial.

The offset axis was resampled at 65 m by simple interpolation before migration, to avoid offset aliasing in the downward continuation. This offset sampling was selected to avoid aliasing artifacts at the reservoir level. Finer sampling would be necessary to migrate shallower events without aliasing the higher frequencies present in the seismic signal. The effects of irregular geometries are more evident at shallow depth, and interesting stratigraphic features are visible in the migrated stack as shallow as 1 km depth. Therefore, future tests may be directed to produce high-quality images shallower than the tests that we present in this paper.

Geometry regularization results

Figure 3 compares the results of geometry regularization of the three methods discussed above for one line. Figure 3a shows the results for normalization, Figure 3b shows the results for

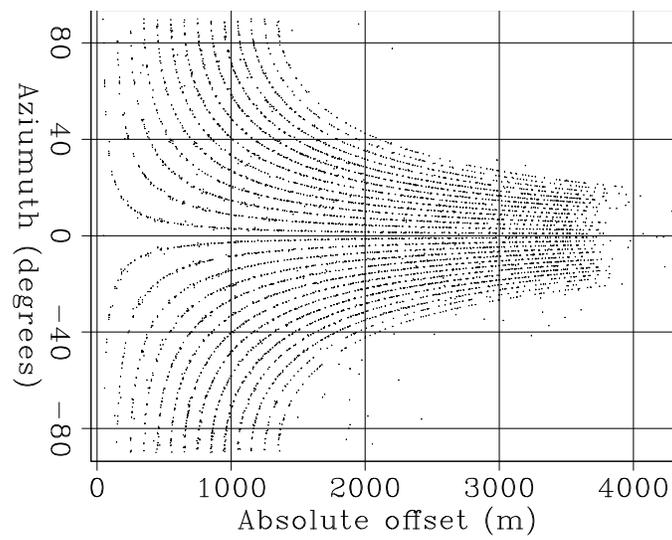


Figure 1: Offset-azimuth distribution of the land data set from South America.
biondo1-OffAz [CR]

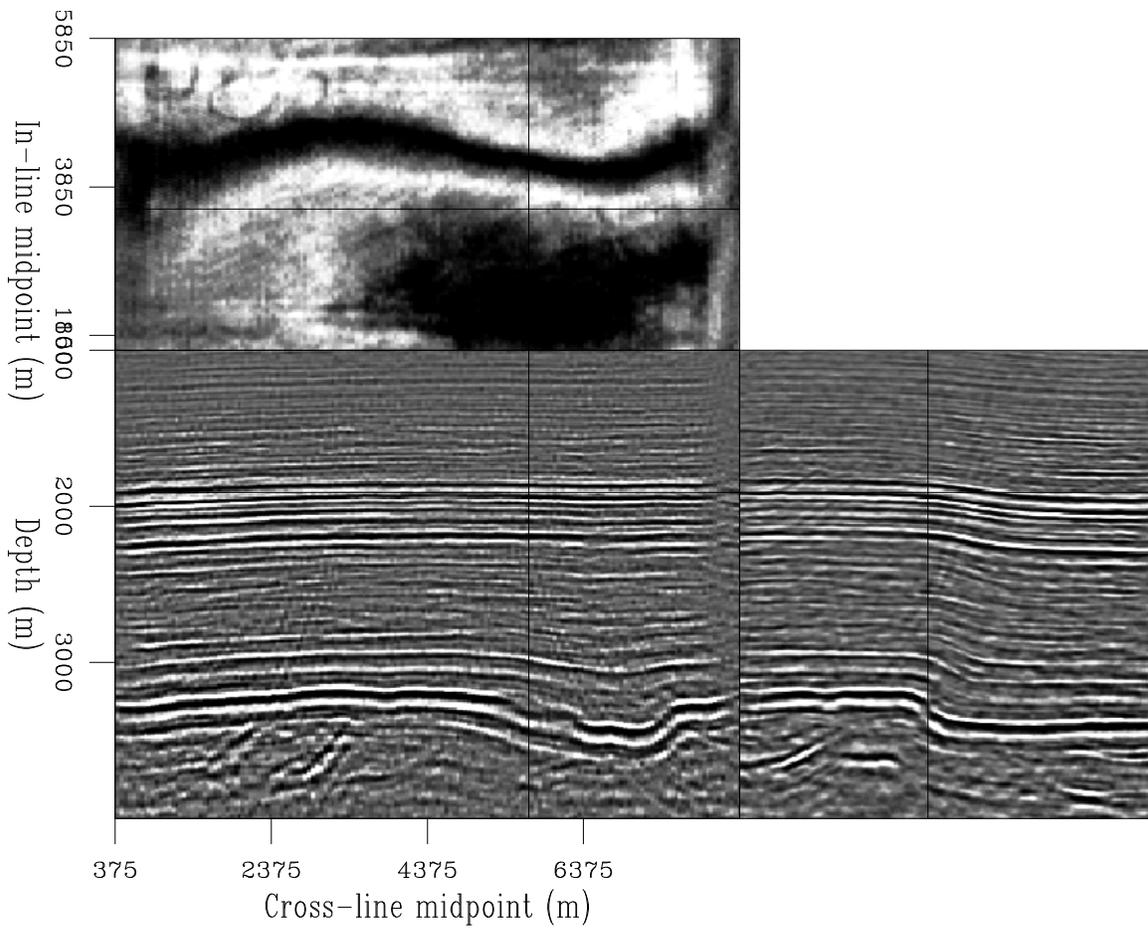


Figure 2: Near-angle stack of the migrated cube.
biondo1-Wm_lin_all_reg_vz_mig_nang_stack [CR]

regularization, and Figure 3c shows the results for AMO regularization. Comparing the in-line sections at one offset (3.38 km) shows the advantages of both regularization and AMO regularization over simple normalization. The data amplitudes after normalization are still fairly uneven, and thus they are likely to produce artifacts during the migration. On the contrary, in the data obtained using regularization the data amplitudes are better balanced. The steeply dipping reflection from the fold at the reservoir level is better preserved in the AMO regularization results than in the simple regularization results. The reason is quite apparent when examining the data as a function of offset for one particular midpoint location. The dipping event is smiling upward after NMO, and thus it is attenuated by simple smoothing over offset.

Figure 4 shows a detail of the same line that illustrates the effects of the regularization term during the process. As for Figure 3, Figure 4a shows the results for normalization, Figure 4b shows the results for regularization, and Figure 4c shows the results for AMO regularization. An acquisition gap is clearly visible in the middle of the constant-offset (2.275 km) section in panel (a). Simple normalization cannot fill the gap. On the contrary, the gap is filled in the regularized results, that exploit the information from the neighboring offsets. The gap in the dipping event is better filled by the AMO regularization because the information from neighboring offsets is moved to the missing data consistently with their kinematics. The differences in behavior between the two regularization methods are apparent when analyzing the time slices in the upper part of the figure. As the dipping event moves towards higher offset (up), it also moves towards further midpoint (right). This movement is well seconded by the AMO regularization, while it is smoothed over by the simpler regularization scheme.

Amplitude-preserving migration results

We migrated the data after geometry regularization using common-azimuth migration and produced three different prestack migrated images with the Common Image Gathers (CIG) function of the in-line offset ray parameter (p_{hx}). Figure 5 compares one line extracted at $CMPY=5.775$ km, for a wide reflection angle. As before, panel (a) displays the result obtained using normalization, panel (b) displays the result obtained using regularization, and panel (c) displays the result obtained using AMO regularization. The image obtained by normalization is more noisy than the ones obtained by using regularization. The fold is slightly better imaged in the results using AMO regularization (panel c). These differences are confirmed in the depth slices shown in Figure 6 shows two depth slices taken at the reservoir level (depth 3.27 km) obtained by migration after geometry regularization with: normalization (top), regularization (middle) AMO regularization (bottom). Panels (a) show the images for a narrow reflection angle, panels (b) show the images for a wide reflection angle. The depth slices obtained using normalization are noisier than the ones obtained using regularization. The dipping event (marked as Fold in the figures) is better imaged in both the normalization results and in the AMO regularization results than in the regularization results.

There are also other subtle differences between the two images obtained with regularization. Only a joint analysis of the seismic images with the well logs from the existing well could determine if they are significant for the accuracy of reservoir characterization.

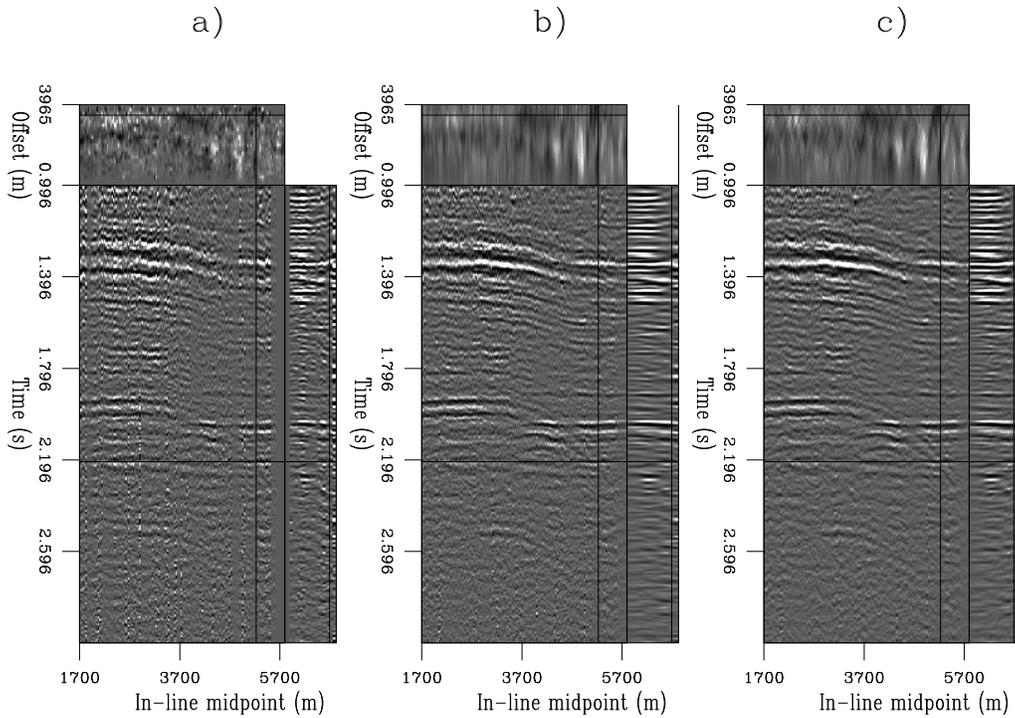


Figure 3: A prestack line after geometry regularization with: a) normalization, b) regularization, c) AMO regularization. `biondo1-Data_abc` [CR]

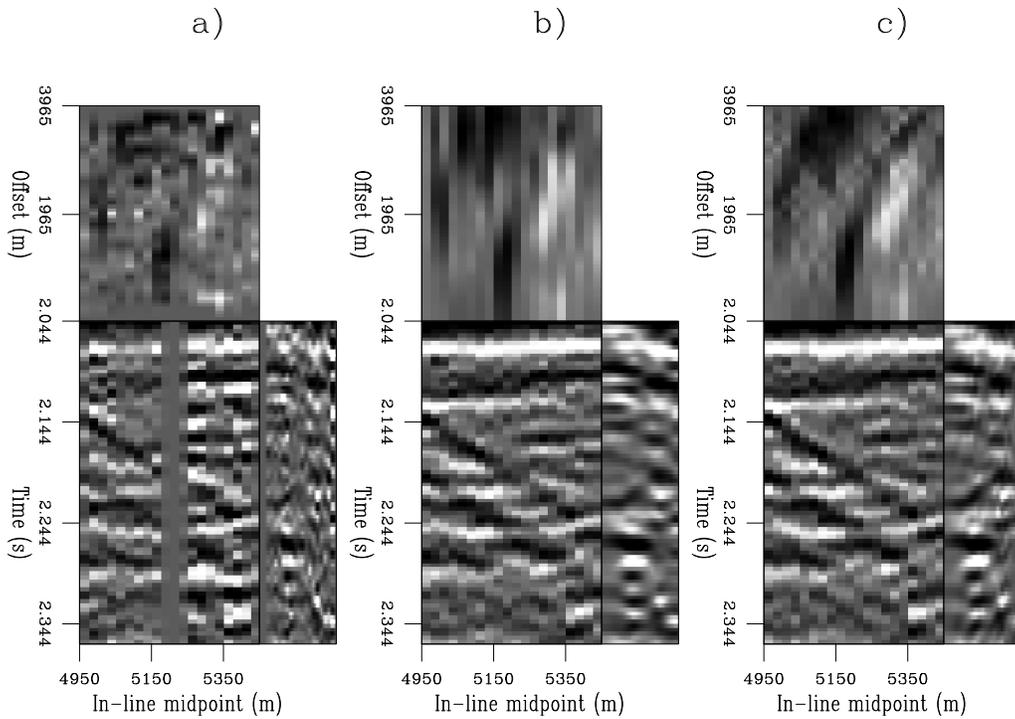


Figure 4: A detail of the prestack line shown in Figure 3. Notice the data gap in the middle of the common-offset (2.275 km) section, and how it has been interpolated differently by regularization (panel b) and AMO regularization (panel c). `biondo1-Data_win_abc` [CR]

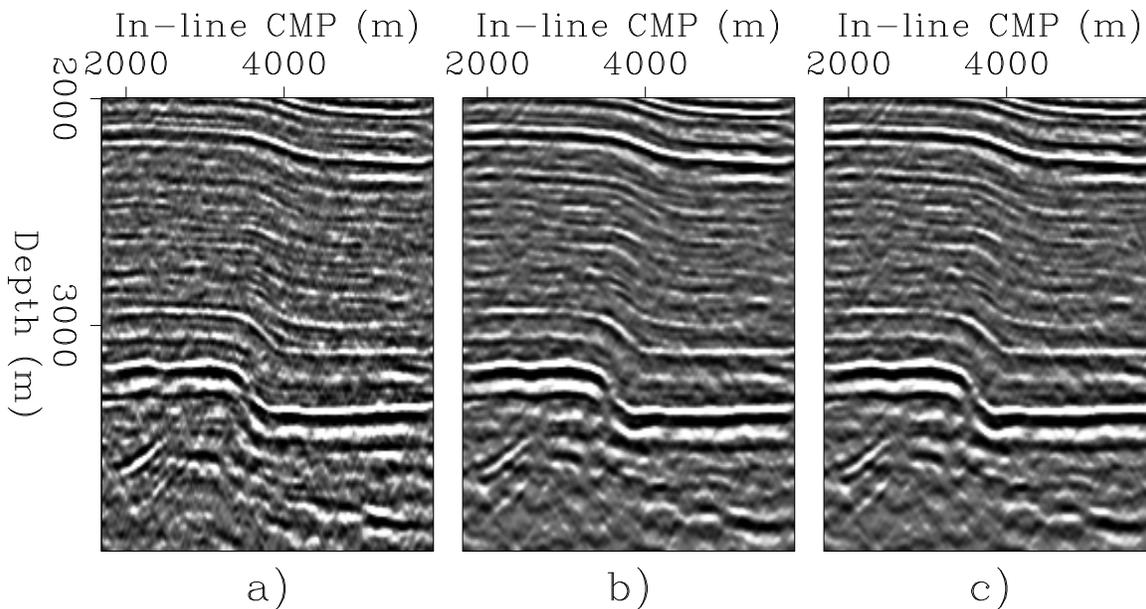


Figure 5: Images for wide reflection angle at constant cross-line location (CMPY=5.775 km). The images were obtained after geometry regularization with: a) normalization, b) regularization, c) AMO regularization. `biondo1-Line_abc` [CR]

Figure 7 shifts our attention to a shallower level, where there are no obvious structural features, but meandering channels are visible. The depth slices were taken at a depth of 1.91 km. One channel is clearly visible in the upper left corner of the images, while another one (less clearly) visible around the lower right corner of the images. As for the previous figures, the slices on the left are taken for a narrow reflection angle, the slices on the right are taken for a wide reflection angle. The images obtained using normalization are noisier and show more clearly the oblique acquisition footprints. In the narrow angle image, the noise is so overwhelming that no channels are visible. The two regularization results are comparable for narrow angles. At wider angle the AMO regularization image is slightly better focused. Subtle differences are noticeable in the imaging of both channels, where indicated by the arrows. Although the images of the upper channel are unfortunately affected by edge effects caused by the migration.

CONCLUSIONS

The proposed method for regularizing the geometry of 3-D prestack data set performed well in a real-data test. The analysis of the data after geometry regularization demonstrates that the regularization methods fill the acquisition gaps by using the information from neighboring offsets/azimuths. Therefore, they provide a better input to the migration than the simple normalization by the partial stack fold. The imaging results confirm this analysis.

The inclusion of the AMO operator in the regularization assures better preservation of the

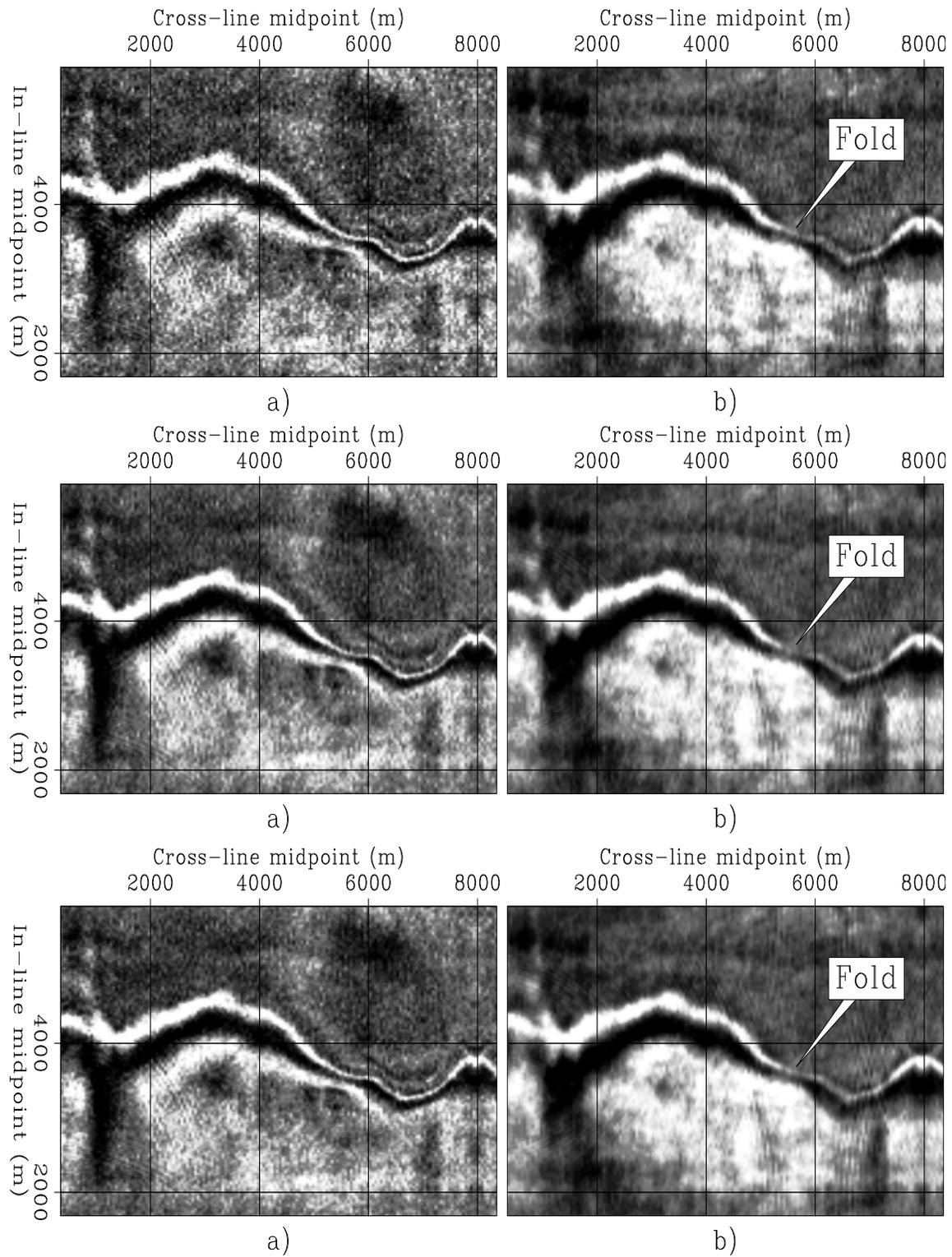


Figure 6: Depth slices, at a depth of 3.27 km, obtained by migration after geometry regularization with: normalization (top), regularization (middle) AMO regularization (bottom). Narrow reflection angle (left), and wide reflection angle (right). `biondo1-3270` [CR]

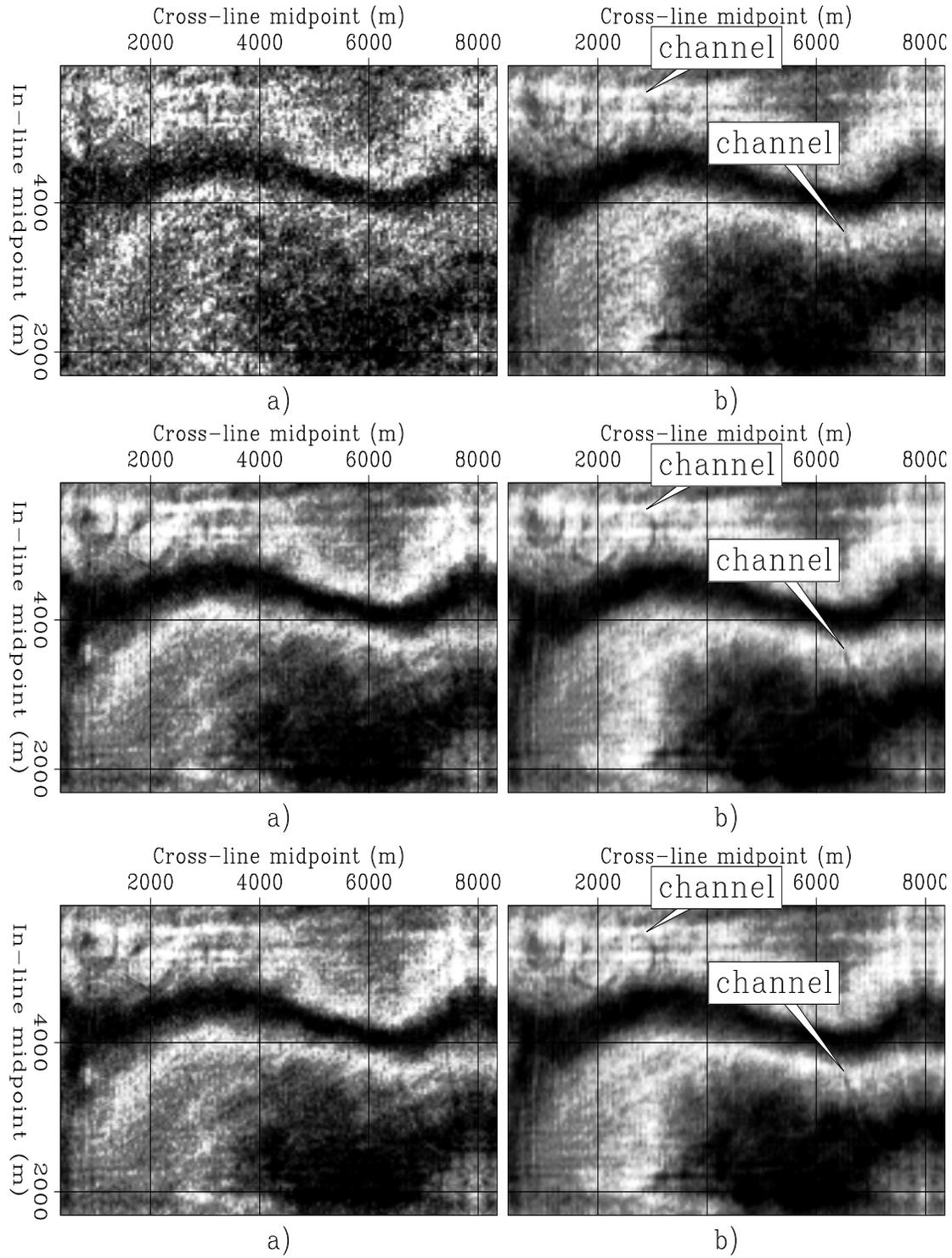


Figure 7: Depth slices, at a depth of 1.91 km, obtained by migration after geometry regularization with: normalization (top), regularization (middle) AMO regularization (bottom). Narrow reflection angle (left), and wide reflection angle (right). `biondo1-1910` [CR]

steeply dipping event, thus yielding higher-resolution images than when the AMO operator is not applied. At the reservoir level (depth of 3.2 km), the improvements in the images are fairly subtle. We would expect more substantial differences for shallower events. However, the processing parameters (in particular, offset sampling for migration) were not optimized for shallow targets. Furthermore, in the tests presented in these paper we ignored the data azimuth. Preliminary tests indicated that taking into account the data azimuth when regularizing the geometry can improve the results at far offset. The method and the codes are ready for these further tests.

An obvious improvement to the method is to allow the the smoothing range parameter (ρ_D) to vary. A non-stationary regularization could better take into account the local sparseness in the data. Using ρ_D that changes laterally could be dangerous and lead to instability, but making it function of depth should be trivial and makes sense because the smoothing range ought to be constant as a function of the reflection angle at depth, not of the offset at the surface.

ACKNOWLEDGMENTS

We would like to thank Trino Salinas from Ecopetrol for bringing the data set to SEP during his visit this summer, and Ecopetrol for making the data available to SEP.

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., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: *Geophysics*, **63**, no. 2, 574–588.
- Biondi, B. L., 1999, 3-D Seismic Imaging: <http://sepwww.stanford.edu/sep/biondo/Lectures/index.html>.
- 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., and Biondi, B., 1997, Equalization of irregular data by iterative inversion: Expanded Abstracts, 1115–1118.
- Chemingui, N., 1999, Imaging irregularly sampled 3D prestacked data: SEP-101.
- Claerbout, J., and Nichols, D., 1994, Spectral preconditioning: SEP-82, 183–186.

- Claerbout, J., 1998, Geophysical estimation by example: Environmental soundings image enhancement: <http://sepwww.stanford.edu/sep/prof/>.
- 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.
- Fomel, S., 2001, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.
- Nemeth, T., Wu, C., and Schuster, G. T., 1999, Least-squares migration of incomplete reflection data: *Geophysics*, **64**, no. 1, 208–221.
- 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., 2001, Model-space vs data-space normalization for finite-frequency depth migration: *SEP-108*, 81–90.
- Ronen, S., 1987, Wave-equation trace interpolation: *Geophysics*, **52**, no. 7, 973–984.
- 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.
- Sava, P., and Biondi, B., 2001, Amplitude-preserved wave-equation migration: *SEP-108*, 1–26.
- Strang, G., 1986, *Linear Algebra and its Applications*: Harcourt Brace Jovanovich College Publishers.
- Vlad, I., and Biondi, B., 2001, Effective AMO implementation in the log-stretch, frequency-wavenumber domain: *SEP-110*, 63–70.

Wave-equation prestack depth migration for sub-basalt P and converted wave imaging

Morgan Brown, Biondo Biondi, and Clement Kostov¹

ABSTRACT

We apply 2-D wave-equation prestack depth migration to a North Sea basalt dataset. High velocity and impedance contrasts across a basalt layer causes low P -wave reflection signal-to-noise ratio under the basalt. We first migrate the P -wave data with a provided depth velocity model, and then replace sub-basalt P -wave velocity with a simple estimate of shear wave velocity and migrate locally-converted shear waves. Angle domain common image gathers assist our interpretation of a sub-basalt converted wave.

INTRODUCTION

Layers of high-velocity sediments, such as basalt, salt, and carbonates, impede conventional P -wave reflection imaging when the velocity contrast between the high-velocity layer (HVL) and overburden is large. In particular, there are considerable challenges associated with imaging below the flood basalts of the Faeroes-Shetland Basin in the North Sea, though the potential returns generate considerable exploration interest (Ogilvie et al., 2001; Sweetman, 1997). The large velocity and impedance contrasts across the sediment/top-basalt interface seriously hamper traditional time imaging, and decrease available velocity information due to poor coverage in P -wave incidence angle. Based on the quality of current published imaging results over basalt in this region, the key questions on which further investment hinge—basalt thickness, tertiary sediments under basalt, and sub/intra-basalt fault distribution—remain largely unanswered.

For incidence angles greater than the P -wave critical angle, and when the overburden compressional wave velocity is similar to the shear wave velocity in the HVL, a relatively large amount of locally-converted S -wave energy penetrates into the HVL (Ogilvie and Purnell, 1996). In some cases, these locally-converted reflections are energetic enough to be recorded at the further offsets of conventional streamer data. In theory, the “symmetric” converted modes (energy travels as S through the same layers on the source and receiver legs of the raypath) can be imaged by off-the-shelf P -wave algorithms. In practice, however, many factors inhibit the success of sub-basalt converted wave imaging. Strong multiple reflection and refracted waves with complex raypaths (Longshaw et al., 1998) and low signal-to-noise ratio (Hanssen et al., 2000) render identification of converted modes in the prestack domain

¹email: morgan@sep.stanford.edu, biondo@sep.stanford.edu, clement.kostov@westerngeco.com

difficult. Additionally, the basalt flows often exhibit a highly complex structure, both laterally (rugose) and vertically (multiple flows, thin layering), the result of which is degradation of the coherency of converted modes (Lafond *et al.*, 1999).

In this paper, we apply 2-D prestack wave equation depth migration to long-offset streamer data recorded in the Faeroes-Shetland Basin by WesternGeco and discussed previously by Kostov *et al.* (2000). The migration program produces angle domain common image gathers (ADCIG) (Prucha *et al.*, 1999; Sava and Fomel, 2000). Because it is parametrized in terms of an actual physical parameter, we show that the angle domain clearly illustrates critical angle as a function of depth, and is thus well-suited to the task of discriminating between converted wave reflections and other modes. After making appropriate substitutions for S -wave velocity in a depth velocity model, we migrate the symmetric and asymmetric converted wave modes in the data and interpret an event on the PSSP section as the base basalt reflection. We then discuss the evidence in support of that interpretation.

DATA BACKGROUND

The data were recorded by WesternGeco in 1999, as part of a multi-client (non-exclusive) survey which acquired a coarse grid of ten 2-D lines. A single 2-D line, consisting of 2943 shots covering roughly 120 km, was donated to SEP. The survey used the following acquisition parameters:

- Shot interval: 25 m
- Group interval: 12.5 m
- Number of channels: 912
- Active array length: 11.4 km
- Near offset: 185 m
- Record length: 9.8 sec at 4 msec

Of the 2943 total shots, we chose to limit our experimentation to the 960 from the center of the survey that were processed by WesternGeco using the Delft Surface Related Multiple Elimination (SRME) technique (Verschuur *et al.*, 1992). Before demultiple, the subset underwent the following preprocessing:

- Swell noise removal.
- Source signature deconvolution using nominal signature.
- Bandpass filtering and subsampling to 8 msec in time and 25 m group interval.
- Truncation to 9 km maximum offset.

After sorting into CMP gathers with 353 traces, the study area covered 26 km of midpoints. Cable feathering over the 120 km line was variable, but unfortunately quite strong over the subset that we studied. Differences between actual and nominal positions reach 2 km at 9 km offset. The SRME demultiple technique assumes nominal geometry, so the feathering issue in this case likely was a first order contributor to the incomplete suppression of multiples at far offsets (Kostov et al., 2000). As we explain later, success of multiple suppression over the entire offset range is a crucial prerequisite to the successful identification of converted waves in both the prestack data and migrated images, since converted waves and multiples have similar apparent velocities. Other factors, such as 3-D effects and source/receiver directivity, also contribute to poorer performance of the SRME method.

WesternGeco also provided a depth velocity model, shown in Figure 1. The basalt layer exhibits a complicated vertical structure. A thin “transition zone” (perhaps interbedded basalt-sediment layers) with interpreted velocity of 3000 m/s gives way to more basalt-like velocities near 4500 m/s. The base of basalt is interpreted as a relatively flat reflector at roughly 3000 meters depth, underlain by a basement assigned a velocity of 5000 m/s.

MIGRATION ALGORITHM

We use a 2-D Extended Split-Step prestack depth migration algorithm (Stoffa et al., 1990) with three so-called “reference velocities” to handle lateral velocity variation. Image sampling in depth is 16 meters, which is probably a bit coarse for the regions of the image above the basalt, but adequate for sub-basalt imaging.

The migration algorithm outputs angle domain common image gathers (ADCIG) for velocity analysis. In fact, the “angle” axis here is parameterized in terms of offset ray parameter, p_h , which is linked to reflection opening angle θ by the following relationship (Prucha et al., 1999).

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

ϕ is the reflector dip angle, and $V(z, \mathbf{m})$ is the interval velocity in the neighborhood above the reflector point. Sava and Fomel (2000) show an efficient method for direct computation of ADCIGs as a function of opening angle. In both cases, the ADCIGs are easier to interpret for angle-dependent phenomena than are offset-domain common image gathers, a fact that we feel will improve our ability to identify converted waves in the presence of complex multiples and refracted waves.

We summarize some important parameters for the migration below:

- Input data CMP spacing: binned to 25 m (from nominal of 12.5 m).
- Image CMP sampling: 1024 samples @ 25 m (26 km total).
- Image Depth sampling: 325 samples @ 16 m (5.2 km total).
- Frequencies imaged: 3-34 Hz.

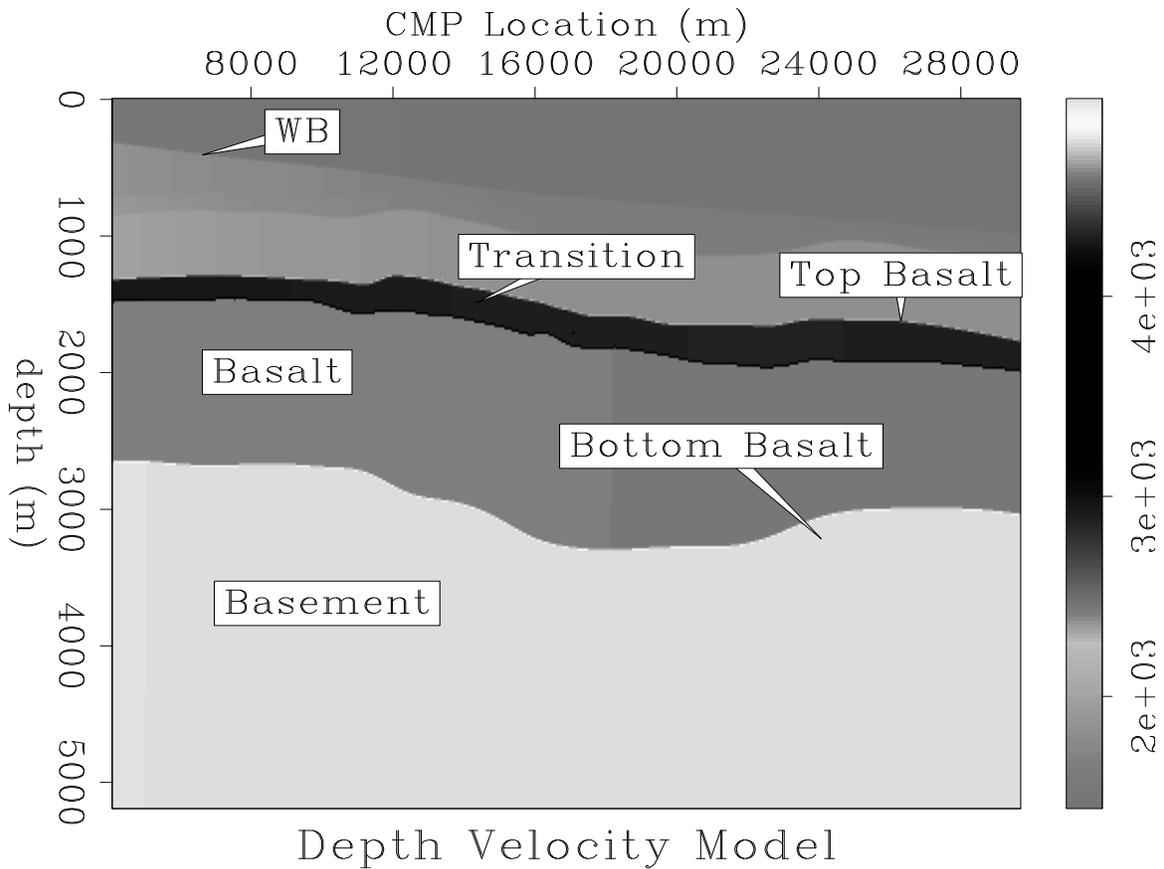


Figure 1: Depth migration velocity model provided by WesternGeco. We used the non-intuitive color scale shown because with the large velocity range (1400-5000m/s), it is impossible to resolve the layer boundaries on a simple grayscale intensity paper plot. Water depth ranges from 300-800 meters. The strongest reflection among the near surface sediments comes from the flat reflector at roughly 1000 meters depth. The transition zone consists of a package of reflectors from the basalt top. Base basalt interpretation is questionable, due to very low interpretable reflection energy. `morgan2-vel-p` [ER]

- ADCIG sampling: 120 samples @ 8×10^{-6} sec/m.

EVENT IDENTIFICATION

In theory, the maximum transmission of locally converted S -wave energy across a HVL occurs at P -wave critical angle. The reflected P -wave also exhibits an amplitude “bloom” at its critical angle. While this P -wave bloom should be readily visible on ADCIGs, Ogilvie and Purnell (1996) show that it is not visible as a function of offset. We observe this amplitude bloom on depth slices of a constant p_h migration cube, and use it as an indicator of critical angle in our search for converted waves. We also believe that overturned, or diving waves may be present on the ADCIGs as events with a high apparent velocity. Fliedner and White (2001) show the utility of diving waves in estimating the thickness and vertical velocity gradient of the basalt.

When identifying converted waves on ADCIGs, we expect to see events with a lower apparent velocity than primary reflections, and which occur only at larger p_h . Multiples will be seen for all p_h (and similarly have a lower apparent velocity than primaries), while primaries will appear below basalt for small p_h only. If the multiple suppression is incomplete at far offsets, we lose the ability to discriminate between multiples and converted waves via their respective amplitude-versus- p_h dependence. As mentioned earlier, this 2-D survey suffered from strong cable feathering, and we have reason to expect that the SRME demultiple method may not fully attenuate multiples at far offsets. For these reasons, we feel that the raw data should be migrated at some point in the future.

RESULTS

We ran four prestack depth migrations, one for four possible permutations of converted waves. For the “PPPP” migration, we image the usual P -wave reflections, using the velocity of Figure 1. In the “PSSP” migration, we image waves which convert from P to S at the top of basalt, reflect as S , then convert back to P at the top of basalt. The PSSP waves are “symmetric”, meaning that by simply replacing compressional wave velocity with an estimate of shear wave velocity below the basalt, we can image these waves with conventional methods (Gratwick, 2000). We simply scaled the velocity model shown in Figure 1 by 0.6 (corresponding to a V_p/V_s ratio of roughly 1.7) below the top basalt. The last two migrations image the “asymmetric” converted waves, “PSPP” and “PPSP”, or waves for which only one leg of the source-to-receiver raypath contains a sub-basalt conversion. Our migration code can accept different velocity functions for the source-to-reflector and reflector-to-receiver legs of the downward continuation.

Since in all migrations, we assume P -wave propagation from source to top of basalt and from top of basalt to the receiver, we datumed the prestack data to 1200 meters, just above the top of basalt.

Figure 2 shows a `Cubep1ot` view of a constant- p_h PPPP migration. The interpreted base

basalt reflection is just below 3000 m at this CMP (see Figure 1). On the p_h gather, we see a strong reflection around 3000 m, although it has quite a bit of curvature, meaning it is imaged with too high a velocity. On a time-migrated section (not shown), we interpret this as a likely internal multiple; one that takes an extra bounce between the top basalt and water bottom (event 2 on Figure 3). The depth slice is taken through the top basalt reflection; notice critical angle behavior around $p_h = 0.0004$. Given the velocity of 2800 m/s just under top basalt, the predicted critical angle ($p_h = 1/2800 \approx 0.00036$) is consistent with the estimate. Thus we can use critical angle as a rough means of validating the velocity model.

PP Migration

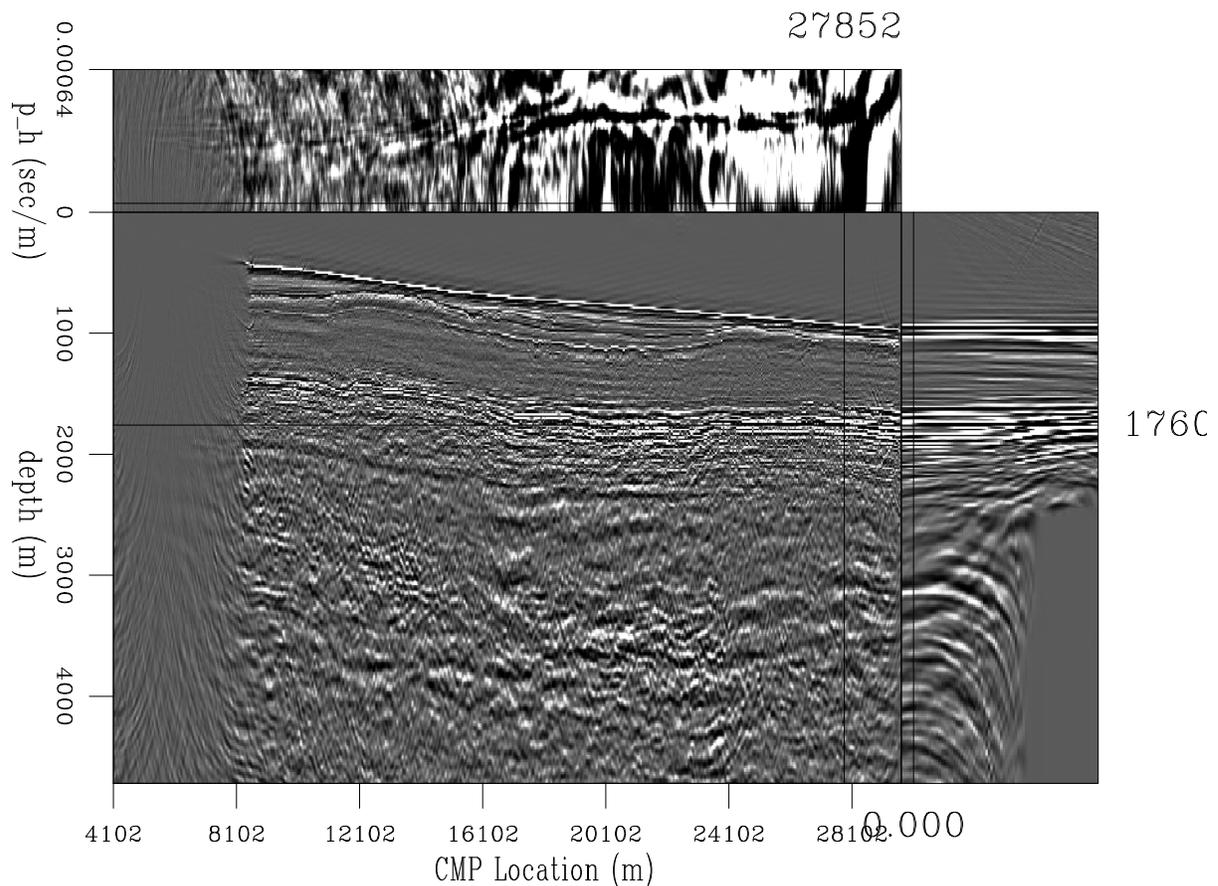


Figure 2: Cubeplot view of PPPP migration. The three axes on the figure are depth, CMP, and p_h . `morgan2-mig-cube` [CR]

Figure 4 shows five selected ADCIGs from the PSSP migration, from CMP 19352 m to 24477 m. The package of strong, low-velocity events is interpreted as the top basalt water layer multiple (one bounce). However, at a depth of about 2700 m on all gathers we notice two persistent, fast events which occurs only at high p_h . The deeper of the two (marked “C”) is fairly flat, while the shallower one (marked “*”) has a higher apparent velocity, and thus curves upward. At CMP 24477, notice a package of strong events (marked “M”) with higher velocity than the obvious multiples below, but still not flat. Initially, we thought these events

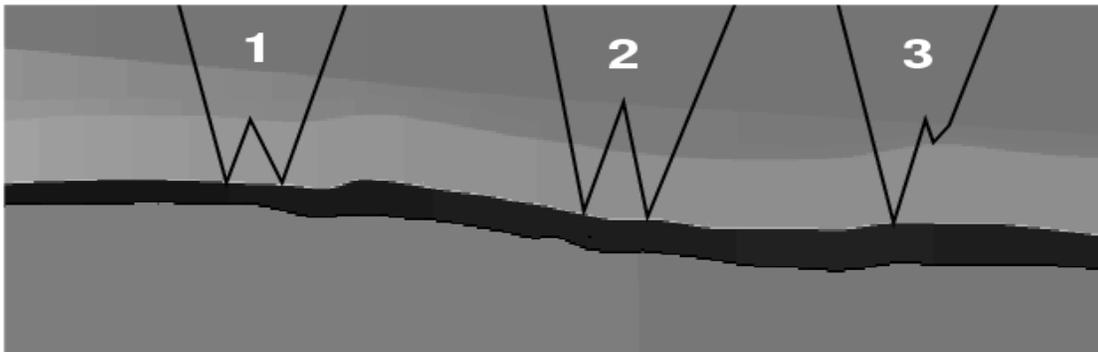


Figure 3: Three possibly energetic internal multiples with a faster apparent velocity than water layer multiples. Because they do not reflect at the free surface, the SRME method is not designed to attenuate them. `morgan2-mult2` [NR]

might be converted waves, but on a time-migrated section, we noted that the events have the same character as top basalt and the correct time delay to be a top basalt/water bottom internal multiple (event 2 on Figure 3). Since the events marked “C” seem to be independent from these probable internal multiples, we are confident in our interpretation of these events as real PSSP reflections.

Figures 5-8 show near-angle stacks (for PPPP) and far-angle stacks (for PSSP, PSPP, and PPSP) of the four constant- p_h migrations. The PPPP migration near-angle ($p_h = [0.0, 0.00016]$) stack (Figure 5) appears to have little interpretable coherent energy below the top basalt. The signal-to-noise ratio appears highest on the right hand side of the image, and we can see a coherent event just below 3000 meters, which has been interpreted as base basalt (compare with Figure 1). A preliminary base basalt has been annotated onto the figure, but the uncertainty is quite high.

The PSSP far-angle ($p_h = [0.00040, 0.000696]$) stack (Figure 6) has more interpretable energy, but much of it is likely due to un-suppressed multiples. The first top basalt water bottom multiple (TBM1) is readily visible, dipping down from 2000 m depth on the left to 4000 m depth on the right. The second multiple (TBM2) is visible near the bottom of the section. On the right side of the image, we have an interesting package of events, which we believe are internal multiples (labeled “IM2”; event 2 on Figure 3), and which appear convex between CMP 24000-28000 m, and then appear to dip down between CMP 20000-24000 m. The possible sub-basalt PSSP event that we interpreted in Figure 4 is annotated with the CMP location (in kilometers) of the gathers.

Time constraints prevented us from more fully interpreting the results of the far-angle ($p_h = [0.00016, 0.000408]$) stacks of the asymmetric mode migrations (Figures 7 and 8). Still a few comparative features merit attention. We note that the first top basalt water bottom multiple focuses differently on each plot, and also a slight relative horizontal shift between the two panels. In perfunctory viewing of the migrated gathers, we saw no obvious candidates for PSPP or PPSP waves. We may not be looking in the right places, or our lack of success may

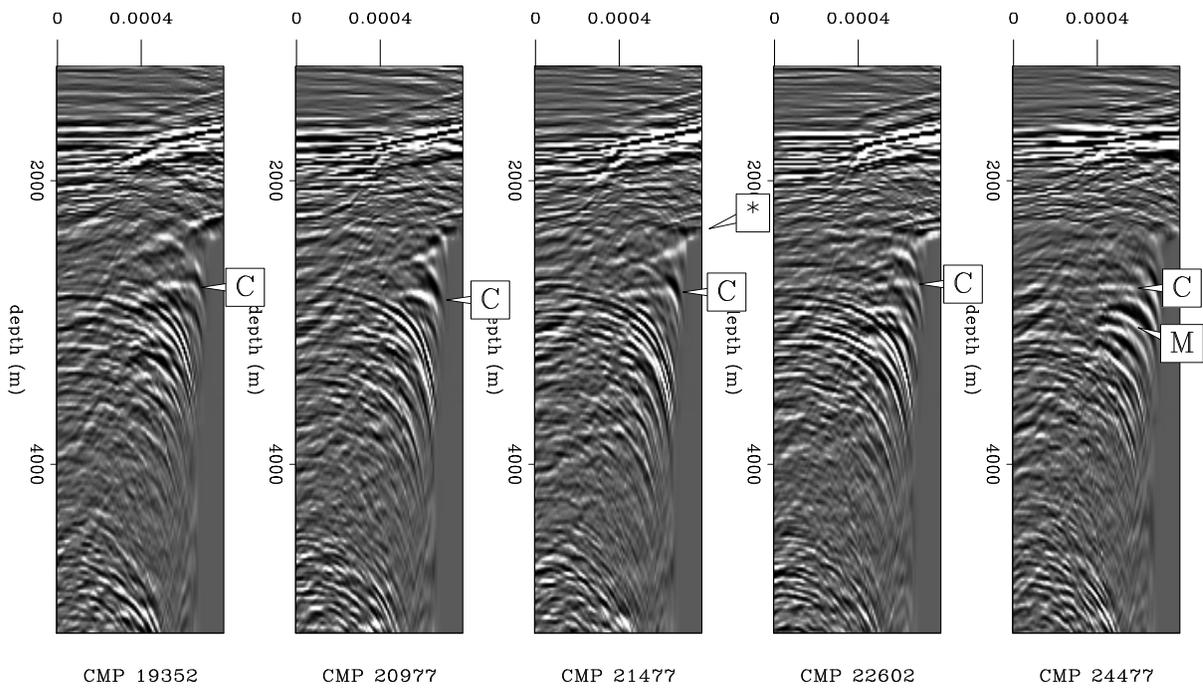


Figure 4: ADCIGs for five CMPs between 19352-24477 m. “C” labels a probable PSSP sub-basalt reflection, possibly the base basalt. “*” denotes a persistent faster event. “M” denotes probable top basalt internal multiple interpreted as base basalt. `morgan2-mig-gathers` [CR]

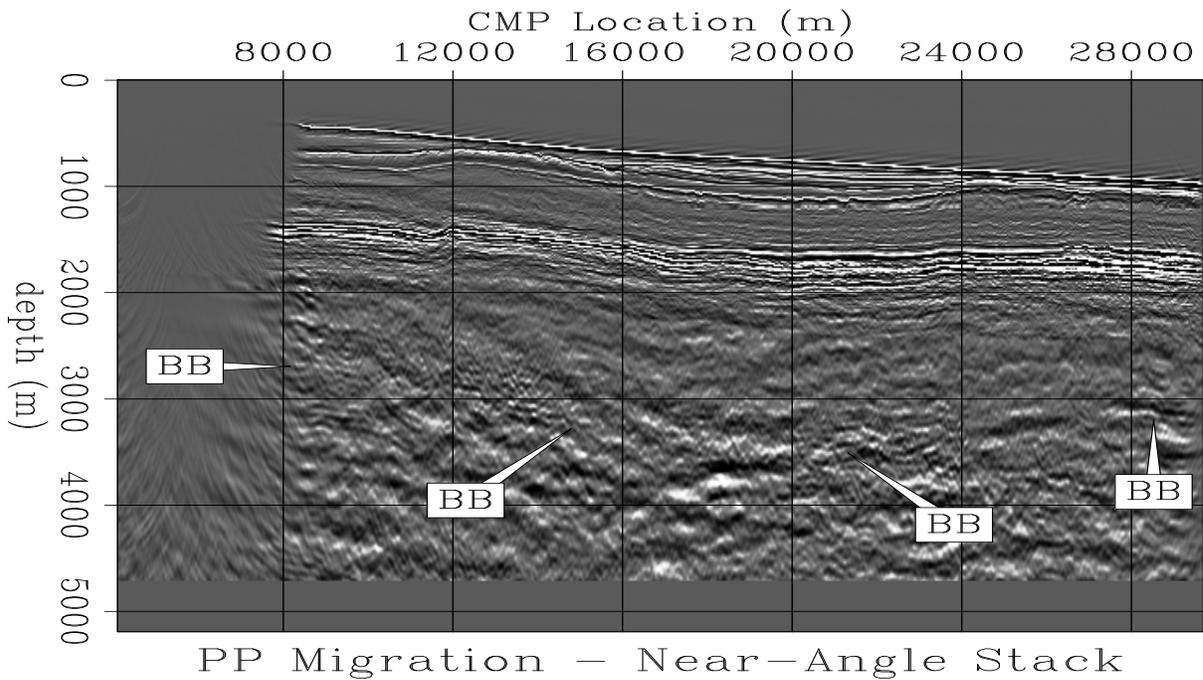


Figure 5: PPP Migration-near-angle stack. Approximate location of interpreted base basalt labeled “BB”. `morgan2-mig-nang` [CR]

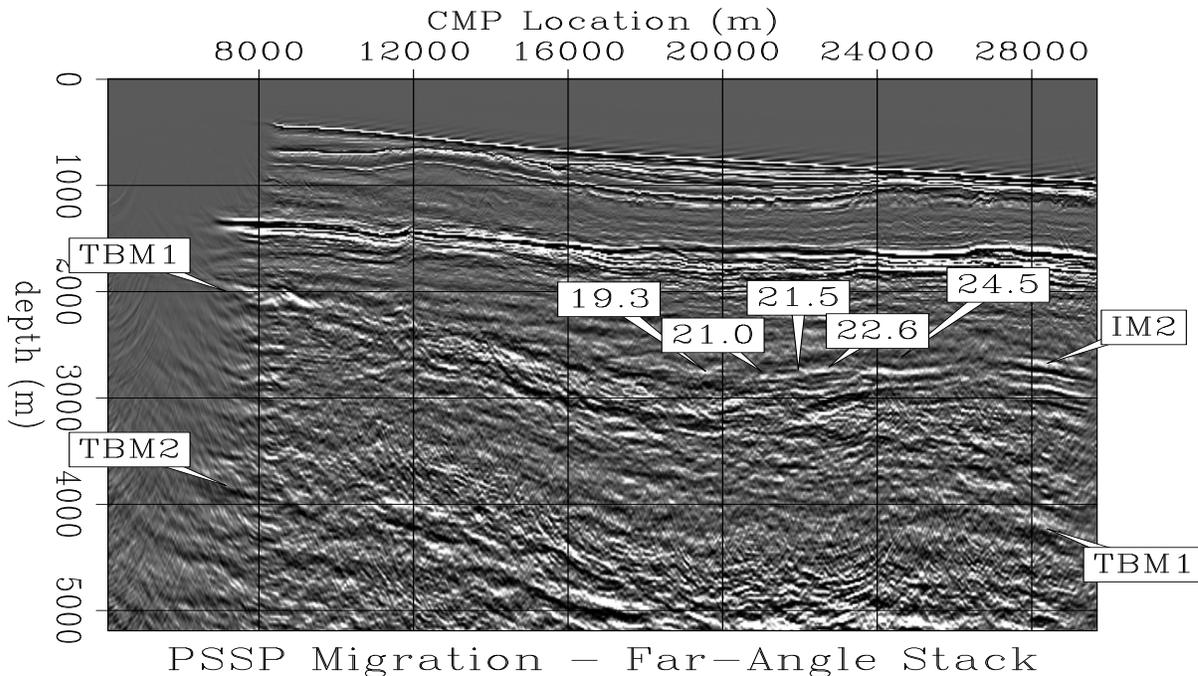


Figure 6: PSSP Migration–far-angle stack. Numeric labels denote CMP locations (in km) of interpreted PSSP waves in Figure 4. “TBM1,TBM2” = first and second top basalt multiples. “IM2” possible internal multiple (event 2 on Figure 3). morgan2-migc-fang [CR]

underscore our need for more serious velocity analysis, or both.

DISCUSSION/FUTURE WORK

We have applied 2-D wave equation prestack depth migration to a long-offset dataset acquired over the flood basalts of the Faeroes-Shetland Basin. We attempted to migrate PSSP, PSPP, and PPSP converted modes by making a suitable replacement of compressional wave velocity with shear wave velocity. In spite of very low subsalt signal-to-noise ratio, we demonstrate that our use of angle domain common image gathers (ADCIGs) facilitated the successful discrimination between a probable PSSP converted wave and multiples in the migrated domain. Still, more investigation and careful processing is required to determine the validity of that interpretation.

Looking to the future, we see many avenues toward improvement. As noted by Kostov et al. (2000) the strong cable feathering in the acquisition of this data contributed to incomplete suppression of multiples at far offsets. Since we used the theoretical amplitude-versus-angle behavior to discriminate between converted waves and multiples, it is crucial either to suppress multiples at all offsets or to image with the raw data.

To improve the performance of the SRME algorithm, we plan to apply azimuth moveout (AMO) (Biondi et al., 1998) to correct for the cable feathering. Although AMO is designed

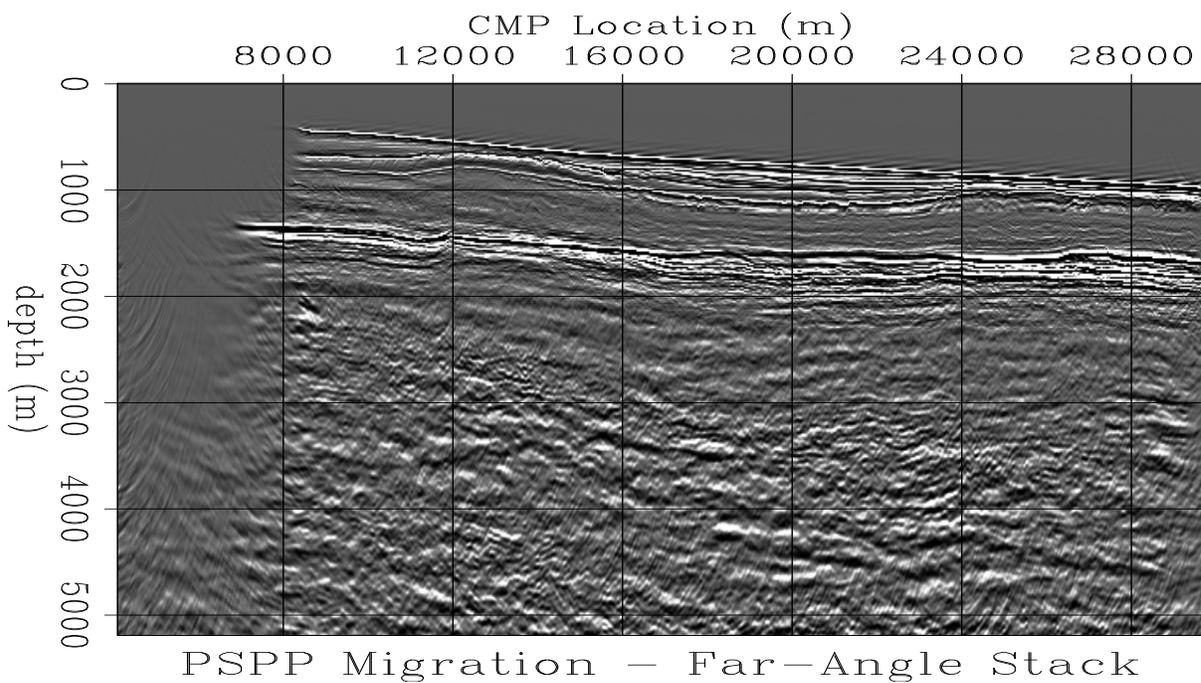


Figure 7: PSPP Migration-far-angle stack. `morgan2-migpspp-fang` [CR]

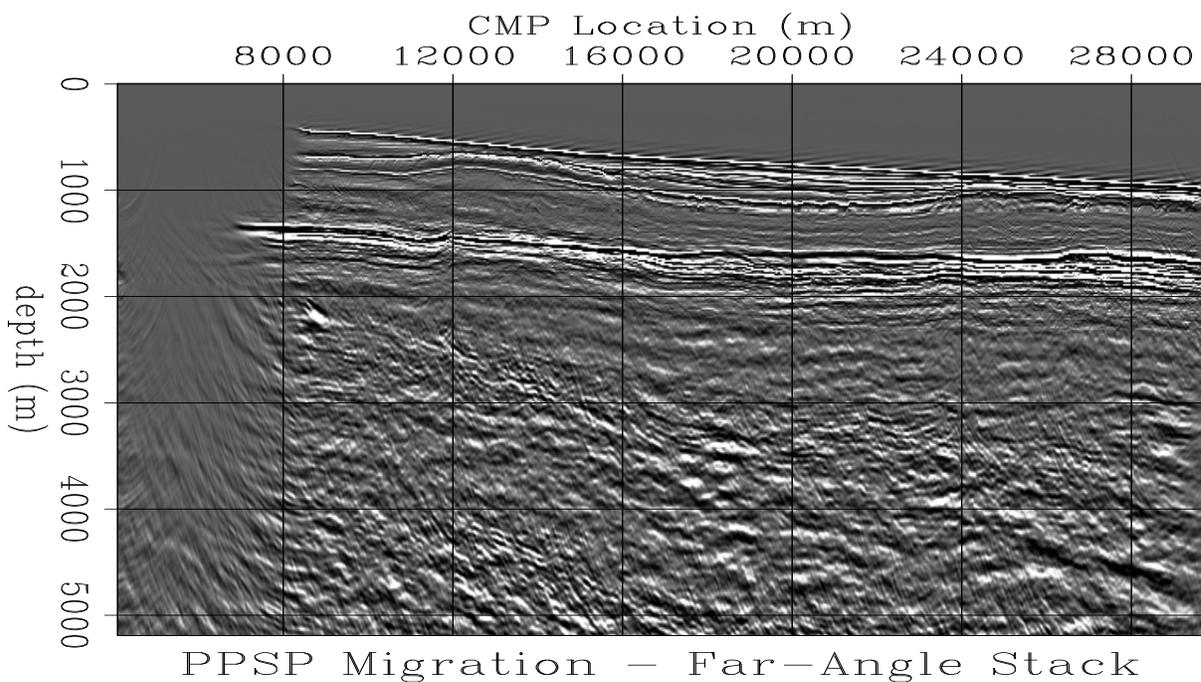


Figure 8: PPSP Migration-far-angle stack. Similar to Figure 7, only shifted slightly laterally. `morgan2-migppsp-fang` [CR]

for use with 3-D acquisition geometries, we can make a first order correction for the feathering by assuming a 2-D earth.

As normally applied in the industry, SRME attenuates multiples from the free surface only. As mentioned in the body of the paper, we noted that the timing of an energetic internal multiple may have been misinterpreted as the base of basalt. At the very least, we need to time-migrate the raw data, as this domain is superior for interpretation of multiples. Next, we believe that careful wave equation or ray-based modeling is crucial to the identification of these more complex multiples. Alternatively (or additionally), if we construct an excellent velocity model of the region above the top basalt, we should be able to migrate the internal multiples.

The strong velocity contrast over the top basalt makes prestack depth migration a must. Unfortunately, the same velocity contrast also seriously inhibits construction of a velocity model, on which depth migration is notoriously dependent. In order to obtain rough bounds on sub-basalt velocity and structure, we may make use of refracted modes. Fliedner and White (2001) demonstrate the use of diving waves (overturning rays), in conjunction with reflection data, to determine basalt thickness and velocity gradient. We also need to investigate the validity of the simplistic substitution we made for shear wave velocity (constant V_p/V_s ratio). While the simplicity of a single-parameter model is appealing, it is likely unrealistic, given the complexity of most basalt flows. Lastly, migration velocity analysis for the asymmetric modes is poorly defined. The recent work of Rosales and Biondi (2001) on Ocean Bottom Cable *PS* data shows promise to this end.

Although we record a single wavefield, the individual modes within that wavefield may be viewed as multiple, possibly independent, sources of data. Lu et al. (2000) demonstrate excellent results from the prestack separation, then depth imaging, of converted waves under salt. Taking this philosophy beyond validation of interpretation of base basalt, we can think of having multiple data sources (prestack separation of converted waves and multiples) which produce a single, self-consistent image (obtained by least squares optimization).

ACKNOWLEDGEMENTS

We would like to thank WesternGeco for its generous donation of the data shown herein, and also for its permission to publish this paper.

Ioan “Nick” Vlad brought to our attention the paper on diving waves. Special thanks to Bob Clapp for his recent contributions to the interpretation software which improved data interpretation greatly (Clapp, 2001).

REFERENCES

Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-D prestack imaging: *Geophysics*, **63**, no. 02, 574–588.

- Clapp, R., 2001, Ricksep: Interactive display of multi-dimensional data: SEP-**110**, 163–172.
- Fliedner, M., and White, R., 2001, Seismic structure of basalt flows from surface seismic data, borehole measurements and synthetic seismogram modeling: Published in Geophysics Online, **May 30, 2001**.
- Gratwick, D., 2000, Subsalt imaging of a 2-D elastic synthetic model, using prestack, split-step, wave equation migration: SEP-**103**, 169–181.
- Hanssen, P., Li, X.-Y., and Ziolkowski, A., 2000, Converted waves for subsalt imaging?: 70th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Session: MC 2.3.
- Kostov, C., Hoare, R., Jaasund, S., and Larssen, B., 2000, Advances in sub-basalt P-wave imaging with long offset streamer data:, *in* 62nd Mtg. Eur. Assn. Geosci. Eng.
- Lafond, C., Kaculini, S., and Martini, F., 1999, The effects of basalt heterogeneities on seismic imaging of deeper reflectors: 69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1433–1436.
- Longshaw, S. K., Sunderland, J., and Horn, I. W., 1998, Mode conversion and multiples: 68th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1340–1342.
- Lu, R., Willen, D., and Watson, I., 2000, Identifying, removing, and imaging P-S conversions at salt-sediment interfaces: 70th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, Session: IMG 1.8.
- Ogilvie, J. S., and Purnell, G. W., 1996, Effects of salt-related mode conversions on subsalt prospecting: Geophysics, **61**, no. 02, 331–348.
- Ogilvie, J. S., Crompton, R., and Hardy, N. M., 2001, Characterization of volcanic units using detailed velocity analysis in the atlantic margin, west of shetlands, united kingdom: TheLeading Edge, **20**, no. 01, 34–50.
- Prucha, M., Biondi, B., and Symes, W., 1999, Angle-domain common image gathers by wave-equation migration: 69th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 824–827.
- Rosales, D., and Biondi, B., 2001, On asymmetric Stolt residual migration for imaging under salt edges: SEP-**110**, 71–74.
- Sava, P., and Fomel, S., 2000, Angle-gathers by Fourier Transform: SEP-**103**, 119–130.
- Stoffa, P. L., Fokkema, J. T., de Luna Freire, R. M., and Kessinger, W. P., 1990, Split-step Fourier migration: Geophysics, **55**, no. 4, 410–421.
- Sweetman, S., 1997, The integration of seismic interpretations with gravity and magnetic 2-D modeling - An example from the offshore Faroes area: The Leading Edge, **16**, no. 01, 15–20.
- Verschuur, D. J., Berkhout, A. J., and Wapenaar, C. P. A., 1992, Adaptive surface-related multiple elimination: Geophysics, **57**, no. 9, 1166–1177.

Time-reversal acoustics and maximum-entropy imaging

James G. Berryman¹

ABSTRACT

Target location is a common problem in acoustical imaging using either passive or active data inversion. Time-reversal methods in acoustics have the important characteristic that they provide a means of determining the eigenfunctions and eigenvalues of the scattering operator for either of these problems. Each eigenfunction may often be approximately associated with an individual scatterer. The resulting decoupling of the scattered field from a collection of targets is a very useful aid to localizing the targets, and suggests a number of imaging and localization algorithms. Two of these are linear subspace methods and maximum-entropy imaging.

INTRODUCTION

Time-reversal acoustics (Fink, 1997; 1999) has become an important research area in recent years because of the many potential applications associated with it and envisioned for it. In biomedical applications, the prototype is the detection and destruction of kidney stones by ultrasound. By insonifying the kidney with pulses of ultrasound and measuring the return signal, it is possible to localize one or more kidney stones and then to send a return signal back to the scatterer with high enough amplitude to cause the stone to fragment. Then, the remaining small pieces eventually pass harmlessly from the system. Training the sound to find the kidney stone automatically is the main purpose of time-reversal acoustics in this application. Analyses of why it works so well and means of improving its performance have been studied by many authors (Prada *et al.*, 1991; 1995; Prada and Fink, 1994; Mast *et al.*, 1997; Devaney, 1999; Blomgren *et al.*, 2001; Tsogka and Papanicolaou, 2001). Most of these analyses have concentrated on imaging of or focusing on small point-like targets.

Chambers and Gantesen (2001) have recently shown that a single spherical acoustic scatterer has from two to four eigenfunctions associated with it. To understand this result, it might help to put it in a larger context by considering scattering from a single spherical elastic scatterer imbedded in an otherwise homogeneous elastic medium. Then, it is well-known that the principal scattering terms arise from changes in bulk modulus (K), density (ρ), and shear modulus (μ). Bulk modulus changes produce monopole scattering; density changes produce dipole scattering; and shear modulus changes produce quadrupole scattering. There is at most

¹email: berryman@sep.stanford.edu

one contribution from monopole scattering; at most three from dipole scattering; and, for acoustics, there is no quadrupole scattering, as there are no viable shear waves. The scattering multiplicity found by Chambers and Gautesen is then understood as, at most $1 + 3 = 4$, while some of the dipole terms may not be excited because of the scatterer-to-sensor array orientation. The monopole term is the one usually treated in analyses of time-reversal acoustics, but for some situations — such as air bubbles in liquid — the dipole terms should also be considered. Nevertheless, we will only consider the monopole contributions in this paper. The generalization of the methods to be presented here in order to incorporate other modes is actually straightforward, as we shall see, and, except for the fact that the trial vectors used for imaging must be tailored to these other modes, the analysis proceeds without any significant alteration.

The fundamental concepts used in time-reversal imaging for acoustics are closely related to concepts in optimal probing and imaging for a variety of physical problems including work by Isaacson (1986), Gisser *et al.* (1990), Cherkaeva and Tripp (1996a,b), Colton and Kirsch (1996), Kirsch (1998), and Brühl *et al.* (2001).

The next section presents a review of the relevant issues in acoustic scattering and time-reversal signal processing. Then I discuss the linear subspace methods of imaging including the well-known MUSIC algorithm, as well as some modifications of MUSIC, and related algorithms. The following section introduces the maximum-entropy imaging approach which is often the preferred method when data are sparse. Then I show some examples and the final section discusses conclusions.

ACOUSTIC SCATTERING AND TIME REVERSAL

Time-reversal acoustics can be studied either in the time domain or in the frequency domain. The two domains have complementary advantages and disadvantages.

Pluses and minuses for the time domain include: Acoustical experiments are generally performed in the time domain. Self-focusing in any acoustic medium can be achieved using time-reversal processing in the time domain. Delay-time windowing easily distinguishes arrivals from well-separated scatterers in the time domain. But, if we want to understand the eigenvector/eigenvalue structure of the acoustic scattering operator in the time domain, we can only obtain one eigenvector at a time, and to obtain more eigenvectors (after the first one) a rigorous orthogonalization procedure must be followed.

Advantages of the frequency domain include the fact that eigenvector/eigenvalue or singular value decomposition (SVD) is quite easily done there, although it is done then one frequency at a time. Thus, all eigenvalues/eigenvectors or singular values/singular vectors are obtained simultaneously. Linear subspace imaging methods are then also most easily applied in the frequency domain. The main disadvantage of the frequency domain is that it is not simple to take advantage of target separation in space (and therefore in time) in this domain.

We will return to discuss more about time domain processing later in the paper, but for now we assume that all measured time traces have been Fourier transformed into the frequency

domain, and that the pertinent equation to study is therefore the Helmholtz equation.

Acoustic wave scattering

We assume that the problems of interest are well-approximated by the inhomogeneous Helmholtz equation

$$[\nabla^2 + k_0^2 n^2(\mathbf{x})] u(\mathbf{x}) = s(\mathbf{x}), \quad (1)$$

where $u(\mathbf{x})$ is the wave amplitude, $s(\mathbf{x})$ is a localized source function, $k_0 = \omega/c_0 = 2\pi f/c_0 = 2\pi/\lambda$ is the wavenumber of the homogeneous background, with ω being angular frequency, f frequency, c_0 the assumed homogeneous background wave speed, and λ wavelength. The acoustic index of refraction is

$$n(\mathbf{x}) = \frac{c_0}{c(\mathbf{x})}, \quad (2)$$

where $c(\mathbf{x})$ is the wave speed at spatial location \mathbf{x} . Thus, $n^2(\mathbf{x}) = 1$ in the background and $a(\mathbf{x}) = n^2 - 1$ measures the change in the wave speed at the scatterers.

Pertinent fundamental solutions for this problem satisfy:

$$[\nabla^2 + k_0^2] G_0(\mathbf{x}, \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}') \quad (3)$$

and

$$[\nabla^2 + k_0^2 n^2(\mathbf{x})] G(\mathbf{x}, \mathbf{x}') = -\delta(\mathbf{x} - \mathbf{x}'), \quad (4)$$

for the homogeneous and inhomogeneous media, respectively. In both cases, we assume the radiation (out-going) boundary condition at infinity.

The well-known solution of (3) for the homogeneous medium in 3D is

$$G_0(\mathbf{x}, \mathbf{x}') = \frac{e^{ik_0|\mathbf{x}-\mathbf{x}'|}}{4\pi|\mathbf{x}-\mathbf{x}'|}. \quad (5)$$

The fundamental solution of (4) for the inhomogeneous medium can be written in terms of that for the homogeneous one (in the usual way) as

$$G(\mathbf{x}, \mathbf{x}') = G_0(\mathbf{x}, \mathbf{x}') + k_0^2 \int a(\mathbf{y}) G_0(\mathbf{x}, \mathbf{y}') G(\mathbf{y}, \mathbf{x}') d^3 y. \quad (6)$$

Note that the right hand side depends on the values of the G , which is to be determined by the same equation. So this is an implicit integral equation that must be solved for $G(\mathbf{x}, \mathbf{x}')$. The regions of nonzero $a(\mathbf{y})$ are assumed to be finite in number (N), in compact domains Ω_n , all of which are small compared to the wavelength λ . Then, there will be some position \mathbf{y}_n (certainly for convex domains) inside each domain Ω_n , characterizing the location of each of the N scatterers. We also call these scatterers “targets,” since it is their locations that we seek.

With these assumptions, it is a good approximation to set the \mathbf{y} arguments of G_0 and G inside the integral equal to \mathbf{y}_n for all \mathbf{y} 's inside domain Ω_n . Then, the fundamental solutions can be moved outside of the integral. There remains the integral over $a(\mathbf{y})$, incorporated into the scattering coefficient

$$q_n \equiv k_0^2 \int_{\Omega_n} a(\mathbf{y}) d^3 y, \quad (7)$$

which then characterizes the strength of the scattering from the n th target domain.

With these definitions, we finally have

$$G(\mathbf{x}, \mathbf{x}') \simeq G_0(\mathbf{x}, \mathbf{x}') + \sum_{n=1}^N q_n G_0(\mathbf{x}, \mathbf{y}_n) G(\mathbf{y}_n, \mathbf{x}'). \quad (8)$$

Furthermore, if the scatterers are sufficiently far apart and the scattering strengths q_n are not too large, then $G(\mathbf{y}_n, \mathbf{x}')$ on the far right can be replaced by $G_0(\mathbf{y}_n, \mathbf{x}')$, giving the explicit formula

$$G(\mathbf{x}, \mathbf{x}') \simeq G_0(\mathbf{x}, \mathbf{x}') + \sum_{n=1}^N q_n G_0(\mathbf{x}, \mathbf{y}_n) G_0(\mathbf{y}_n, \mathbf{x}'). \quad (9)$$

Equation (9) is the Born approximation to $G(\mathbf{x}, \mathbf{x}')$ for small scatterers.

Acoustic time-reversal data analysis

Time-reversal acoustics can be understood in a straightforward way from the scattering theory presented so far. First, define a complex vector \mathbf{H}_n relating scattering at point \mathbf{y}_n with all M of the acoustic sensors in a sensor array located at positions $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$. Then,

$$\mathbf{H}_n^T = (G_0(\mathbf{y}_n, \mathbf{x}_1), \dots, G_0(\mathbf{y}_n, \mathbf{x}_M)). \quad (10)$$

By analogy to (10), we define a general vector of the same form \mathbf{H}_r , indexed by a general location in the model space \mathbf{r} , such that

$$\mathbf{H}_r^T = (G_0(\mathbf{r}, \mathbf{x}_1), \dots, G_0(\mathbf{r}, \mathbf{x}_M)). \quad (11)$$

We will term \mathbf{H}_r a “trial vector” at \mathbf{r} , and \mathbf{H}_n one of the set of N “solution vectors.”

[Note that, for inhomogeneous media, if we use G instead of G_0 in (10) and (11), then we have the method known as “matched field processing.” Similarly, if we replace G_0 by the appropriate fundamental solution for the dipole (instead of the monopole) term, then the analysis proceeds again essentially as follows, but the dipole modes can then be used for imaging.]

With these definitions, the fundamental solution in the Born approximation can be rewritten for $m, m' = 1, \dots, M$ as

$$G(\mathbf{x}_m, \mathbf{x}_{m'}) \simeq G_0(\mathbf{x}_m, \mathbf{x}_{m'}) + K(\mathbf{x}_m, \mathbf{x}_{m'}), \quad (12)$$

where the “response matrix” (or transfer matrix)

$$\mathbf{K} = \sum_{n=1}^N q_n \mathbf{H}_n \mathbf{H}_n^T. \quad (13)$$

Elements of the matrix \mathbf{K} are given by

$$K_{m,m'} = K(\mathbf{x}_m, \mathbf{x}_{m'}) = \sum_{n=1}^N q_n G_0(\mathbf{x}_m, \mathbf{y}_n) G_0(\mathbf{y}_n, \mathbf{x}_{m'}) \quad (14)$$

Clearly, the response matrix \mathbf{K} is complex and symmetric.

IMAGING AND INVERSION USING LINEAR SUBSPACE METHODS

One class of imaging methods available for time-reversal imaging of small targets may be called linear subspace methods, of which the best known method is probably MUSIC (Schmidt, 1979; Marple, 1987; Stoica and Nehoral, 1990; Xu and Kaveh, 1996; Stoica and Moses, 1997). The term MUSIC stands for Multiple Signal Classification scheme. The method determines whether or not each vector in a set of vectors is fully or only partially in the range of an operator. If $\mathbf{T} = \mathbf{K}\mathbf{K}^*$ is the operator of interest (*i.e.*, the time-reversal operator), and the complete set of eigenvectors in the range of the operator (*i.e.*, having nonzero eigenvalue) is given by $\{\mathbf{V}_n\}$, then we can choose a test vector \mathbf{H}_r [see (11)] and define the square of the direction cosine between \mathbf{H}_r and any one eigenvector \mathbf{V}_n to be

$$\cos^2(\mathbf{V}_n, \mathbf{H}_r) = |\mathbf{V}_n^{*T} \cdot \mathbf{H}_r|^2 / |\mathbf{H}_r|^2. \quad (15)$$

We are assuming in this formula that all the eigenvectors \mathbf{V}_n are normalized so $|\mathbf{V}_n| = 1$, while \mathbf{H}_r is not necessarily normalized.

MUSIC and variations

In the present application, \mathbf{r} is a vector ranging over all or some discrete subset of the positions in the model space (usually a set of grid points). Then, there are several functionals we could plot in order to produce an “image” of the scatterers. The most common choice is the MUSIC classification functional

$$\text{csc}^2(\{\mathbf{V}_n\}, \mathbf{H}_r) = \frac{1}{1 - \sum_{n=1}^N \cos^2(\mathbf{V}_n, \mathbf{H}_r)}. \quad (16)$$

Another closely related possibility that has similar characteristics (but does not require normalization of \mathbf{H}_r in some implementations) is

$$\text{cot}^2(\{\mathbf{V}_n\}, \mathbf{H}_r) = \frac{\sum_{n=1}^N \cos^2(\mathbf{V}_n, \mathbf{H}_r)}{1 - \sum_{n=1}^N \cos^2(\mathbf{V}_n, \mathbf{H}_r)}. \quad (17)$$

The interpretation of these functionals as cosecants and cotangents in the subspaces determined by the eigenvectors should now be clear. By plotting these functionals, we find that the targets are located at those points where the denominators approach zero, and therefore in locations where the trial vector is entirely in the range of the scattering operator \mathbf{K} , or equivalently in the range of \mathbf{T} .

Now we can ask the question, how do we make use of these ideas if the data available to us are limited? In particular, it might happen that some of the nonzero eigenvalues are quite small compared to the others, and we do not know whether to include the corresponding eigenvectors in the set $\{\mathbf{V}_n\}$ or not. In this case, we can use a variation on the MUSIC scheme by only considering a subset of the eigenvectors, say $n = 1, \dots, N' \leq N$. In this case, either of the two schemes just described is easily modified by restricting the sums to

$$\csc^2(\{\mathbf{V}_n\}', \mathbf{H}_r) = \frac{1}{1 - \sum_{n=1}^{N'} \cos^2(\mathbf{V}_n, \mathbf{H}_r)}, \quad (18)$$

and

$$\cot^2(\{\mathbf{V}_n\}', \mathbf{H}_r) = \frac{\sum_{n=1}^{N'} \cos^2(\mathbf{V}_n, \mathbf{H}_r)}{1 - \sum_{n=1}^{N'} \cos^2(\mathbf{V}_n, \mathbf{H}_r)}. \quad (19)$$

This approach can then be used to test whether certain eigenvectors are really in the range or not by replotting these functions for different values of N' . The scheme just described could also be used to do crude imaging if only a single eigenvector is known, as might happen if we have used time-reversal processing in the time domain and had found only the first eigenvector. When viewing eigenvectors as measurements, we see that using fewer eigenvectors will result in poorer resolution, as less information is then available to constrain the images.

Extended MUSIC for inversion

MUSIC as described so far is what I would call an “imaging method” in the sense that it is attempting to locate objects, but not to quantify their properties (such as the scattering strength q). We will show now that, with only a little more computational effort, we can also determine the scattering coefficients. The resulting algorithm will then be distinguished as an “inversion method,” rather than imaging.

Our data are again the matrix elements of the response matrix \mathbf{K} . We know that \mathbf{K} can be represented equally in two ways by its scattering expansion and its SVD so that

$$\mathbf{K} = \sum_{n=1}^N q_n \mathbf{H}_n \mathbf{H}_n^T = \sum_{m=1}^M \sigma_m \mathbf{V}_m \mathbf{V}_m^T, \quad (20)$$

recalling that there are M sensors in our array, N scatterers, and that the singular vectors are assumed to be normalized to unity.

One formulation of the inversion problem that is very similar in spirit to MUSIC is therefore to find a matrix

$$\tilde{\mathbf{K}} = \sum_{r=1}^{L^2} \tilde{q}_r \mathbf{H}_r \mathbf{H}_r^T, \quad (21)$$

whose elements agree with those of the measured \mathbf{K} according to some criterion such as least-squares. The sum in (21) is taken over grid points in an $L \times L$ (pixel units) model space in 2D. Generalization to model spaces having other shapes and/or to 3D is straightforward, as there is no use made here of any symmetries in the choice of sensor array arrangement, or of the locations of the test pixels chosen as possible targets and indexed by r . We will want to take advantage of the orthogonality of the singular vectors \mathbf{V}_m . So we suppose that it is possible to solve an equation of the form

$$\sum_{r=1}^{L^2} \tilde{q}_r \mathbf{H}_r \mathbf{H}_r^T \simeq \sum_{m=1}^M \sigma_m \mathbf{V}_m \mathbf{V}_m^T, \quad (22)$$

in the least-squares sense for the coefficients \tilde{q}_r associated with some set of points (pixel or voxel centers) in the model space.

Applying the singular vectors \mathbf{V}_m^* to both the right and left sides of (22), we find

$$\sum_{r=1}^{L^2} \tilde{q}_r (\mathbf{H}_r^T \cdot \mathbf{V}_m^*)^2 = \sigma_m, \quad (23)$$

which can then be rewritten in matrix notation as

$$\begin{pmatrix} (\mathbf{H}_1^T \cdot \mathbf{V}_1^*)^2 & \dots & (\mathbf{H}_{L^2}^T \cdot \mathbf{V}_1^*)^2 \\ (\mathbf{H}_1^T \cdot \mathbf{V}_2^*)^2 & \dots & (\mathbf{H}_{L^2}^T \cdot \mathbf{V}_2^*)^2 \\ \vdots & \ddots & \vdots \\ (\mathbf{H}_1^T \cdot \mathbf{V}_M^*)^2 & \dots & (\mathbf{H}_{L^2}^T \cdot \mathbf{V}_M^*)^2 \end{pmatrix} \begin{pmatrix} \tilde{q}_1 \\ \tilde{q}_2 \\ \vdots \\ \tilde{q}_{L^2} \end{pmatrix} = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_M \end{pmatrix}. \quad (24)$$

We see that estimates of the coefficients can now be found by solving this rather large ($M \times L^2$) and underdetermined least-squares system. This form is similar to various kinds of tomography, with the singular values taking the role of the data, and the singular vectors taking the role of the discrete scanning modality. (For comparison, in traveltime tomography the data are the traveltimes, the matrix elements are the ray-path lengths through the cells, and the model parameters are the slownesses in the cells.) Note that the elements of the present matrix are in general complex, as are the scattering coefficients \tilde{q}_r , while the singular values σ_m are all nonnegative real numbers.

The computational situation can be improved significantly by noting the similarity of the matrix elements to the direction cosines already treated in MUSIC. In fact, if we take the magnitude of each matrix element, we have the square of a direction cosine. Furthermore, if we sum these magnitudes along the columns of the matrix, then we have obtained exactly the quantity

$$|\mathbf{H}_r|^2 \sum_{m=1}^M \cos^2(\mathbf{V}_m, \mathbf{H}_r), \quad (25)$$

which — except for the normalization factor $|\mathbf{H}_r|^2$ — is the sum over all the direction cosines associated with the vector \mathbf{H}_r . This column sum is therefore a measure of what we might call the “coverage” of the pixel centered at \mathbf{r} . If there is no coverage, this column sum is zero and the pixel does not contain a measurable target. If the normalized coverage is close to unity, then this pixel is one that MUSIC will clearly identify as a target. For intermediate values, the situation is somewhat ambiguous, as it is in the normal MUSIC algorithm, but clearly the closer the normalized sum is to unity, the more likely it is that the pixel contains a target.

Now we have a clear pixel selection criterion available, based on these column sums. Thus, we can reduce the size of the inversion problem posed in (24) very quickly by throwing out all pixels that have small column sums, and/or, equivalently, those whose values are $\ll |\mathbf{H}_r|^2$. We can classify the remaining pixels by the magnitudes of the normalized sums, choosing to keep only (say) those pixels having the M largest normalized column sums. Or if there are clearly only $N \ll M$ singular values that are nonzero, then we could reduce the problem still further, and keep only those pixels having the N largest normalized column sums. If we have an $M \times M$ matrix when we have finished this selection process, then we can simply invert the matrix to find the values of the remaining \tilde{q} 's. If on the other hand, we have an $M \times N$ matrix that is not square remaining after this process of elimination, then we will again have to solve the inversion problem, by using overdetermined least-squares.

Another possibility when the singular values have a gradual fall off in magnitude as $m \rightarrow M$, but no precipitous drop to zero, is to multiply (24) on the left by a diagonal matrix whose nonzero elements are the singular values to some power p . Then, the resulting equation is

$$\sum_{r=1}^{L^2} \tilde{q}_r \sigma_m^p (\mathbf{H}_r^T \cdot \mathbf{V}_m^*)^2 = \sigma_m^{p+1}, \quad (26)$$

or, in vector/matrix notation,

$$\begin{pmatrix} \sigma_1^p (\mathbf{H}_1^T \cdot \mathbf{V}_1^*)^2 & \dots & \sigma_1^p (\mathbf{H}_{L^2}^T \cdot \mathbf{V}_1^*)^2 \\ \sigma_2^p (\mathbf{H}_1^T \cdot \mathbf{V}_2^*)^2 & \dots & \sigma_2^p (\mathbf{H}_{L^2}^T \cdot \mathbf{V}_2^*)^2 \\ \vdots & \ddots & \vdots \\ \sigma_M^p (\mathbf{H}_1^T \cdot \mathbf{V}_M^*)^2 & \dots & \sigma_M^p (\mathbf{H}_{L^2}^T \cdot \mathbf{V}_M^*)^2 \end{pmatrix} \begin{pmatrix} \tilde{q}_1 \\ \tilde{q}_2 \\ \vdots \\ \tilde{q}_{L^2} \end{pmatrix} = \begin{pmatrix} \sigma_1^{p+1} \\ \sigma_2^{p+1} \\ \vdots \\ \sigma_M^{p+1} \end{pmatrix}. \quad (27)$$

We can apply a MUSIC-like processing scheme to this matrix (that could then be called “weighted MUSIC,” which is similar to some methods suggested by other authors). This approach permits the singular values to determine in a natural way the cutoff (if any) in the contributions from those singular vectors of negligible importance to the inversion. For example, when $p = 2$, these column sums of the magnitudes of these elements are just $|\mathbf{K}^* \mathbf{H}_r|^2$, which are the matrix elements of the time-reversal operator with each vector \mathbf{H}_r .

MAXIMUM-ENTROPY IMAGING METHODS

Motivation

If we consider the structure of the response matrix \mathbf{K} , whenever the number of targets is much less than the array size so that $N \ll M$, the matrix information is very redundant; the matrix is then rank N , where N is the number of targets to be imaged. Yet, the matrix \mathbf{K} for M sensors in our array is $M \times M$, with $M(M+1)/2$ distinct complex elements. The total number of distinct data present is therefore $M(M+1)$, which is to be used to determine the x, y, z -coordinates in 3D of each target and possibly also its scattering strength q . For complex q , this means we need to find at most $5N$ numbers from $M(M+1)$. In the presence of noise or strong inhomogeneities in the background medium, the redundancy may be needed to resolve several targets. In homogeneous backgrounds, the overdetermined nature of this problem is something we need to consider.

Examples of situations in which limited data are available include: (1) Only one transducer is available. Assuming that the transducer is moveable, then it would be possible to collect data in a “synthetic aperture” mode as is commonly done in radar applications (SAR). (2) Only two transducers are available, but some triangulation is then possible. (3) Only one ping is allowed, but many transducers are available (multistatic case). Here, we can collect only one row of the response matrix. (4) Only the primary eigenvector has been found, as in iterative time-reversal processing in the physical domain, but no attempt has been made to find eigenvectors associated with secondary targets.

How much of the information in the scattering data is really needed to solve the inverse problem in these situations? With limited data, how much data is essential to collect to locate and possibly identify the targets of most interest?

Having posed the general problem, we will not try to answer it completely here. Instead, we will show some examples of what can be done to image with restricted data sets.

Maximum-entropy imaging

One approach to imaging with corrupted data or relatively small data sets is known as Maximum-Entropy Imaging (Gull and Daniell, 1978; Schmidt, 1979; Gull and Skilling, 1984). I will now describe this method briefly, although another method that is also based on the maximum entropy concept is the main focus of this section.

Perhaps the most common use of maximum-entropy imaging arises in astronomy, and especially in radioastronomy (Gull and Daniell, 1978). The goal is to find a map of the distribution of radio brightness across the sky. But the data collected and the methods used can produce a series of such maps, each having different resolutions and noise levels. There does not appear to be any preferred or “best” sky map associated with the data from this point of view. To resolve this ambiguity, we can choose to use an image that satisfies a maximum-entropy objective criterion (to be explained soon). The resulting maximum-entropy image

should not be thought of as the “true” map of the sky in this context, but rather as a map that does not lead to conclusions for which there is either very little or no evidence in the data. The analogy to our problem in acoustics of finding relatively isolated targets in an otherwise homogeneous background is also apparent.

The reader interested in understanding the general context of entropy and its role in statistical mechanics and information theory would do well to read the paper by Jaynes (1957). The history of maximum-entropy in astronomy is also beyond our scope here, but I will point out that Gull and Daniell (1978) introduced this algorithm in part as an alternative to the CLEAN algorithm (Högbom, 1974), which is another popular imaging algorithm (although not as popular now as is the maximum-entropy approach) for use with incomplete and noisy data.

Following Gull and Daniell (1978), we define an intensity m_{ij} at the pixel in the (i, j) position of a test map and let \hat{m}_{kl} be the Fourier transform of m_{ij} . Then, suppose further that the data come to us in the form of measurements \hat{n}_{kl} of the Fourier transform \hat{m}_{kl} . Assuming that these measurements have Gaussian errors, with standard deviations σ_{kl} , the data fitting part of the algorithm is a weighted least-squares error term of the form $\sum_{kl} |\hat{m}_{kl} - \hat{n}_{kl}|^2 / \sigma_{kl}^2$. The objective constraint applied is that the entropy functional $S = -\sum_{ij} m_{ij} \ln m_{ij}$, determined by the nonnegative intensities in the final map, is a maximum. This maximum-entropy objective is natural for imaging a sparsely occupied field because it corresponds to a situation where all the intensities have the same value, thus, providing a sky map with no information in the absence of data. Using these two terms to form an overall objective functional, including a Lagrange multiplier for the data constraints, gives

$$Q(m_{ij}, \lambda) \equiv -\sum_{ij} m_{ij} \ln m_{ij} - \frac{\lambda}{2} \sum_{kl} |\hat{m}_{kl} - \hat{n}_{kl}|^2 / \sigma_{kl}^2. \quad (28)$$

One advantage of the second sum in Q is that it satisfies a χ^2 distribution, and, therefore, the value of this sum can be used as a measure of goodness of the data fit. In particular, the expected value of the sum is equal to the number of terms in the sum. The disadvantage of this sum is that it requires prior knowledge of the standard deviations. For other imaging applications, the second term in (28) could be replaced by an output least-squares functional together with a tolerance value as in Morozov’s discrepancy criterion (Morozov, 1967; 1984; Tikhonov and Arsenin, 1979; Groetsch, 1984; Hanke, 1997; Haber *et al.*, 2000).

We will not pursue this approach further here, but instead introduce another approach that makes use of the maximum-entropy imaging concept.

A new maximum-entropy imaging method

We assume that a set of data has been produced using active sources, that the time series response has been measured as a time series, and that these data have been Fourier transformed to produce the response matrix. At this point, there are two possibilities: (1) We could make direct use of the response matrix, or (2) we could continue processing the data by forming the time-reversal matrix and then computing the eigenvectors of the matrix. A third possibility

that will be seen as a special case of the previous ones is that time-reversal data is collected and a single eigenvector is constructed by iteration on the physical system.

For our present purposes, in any of these cases we can take the diagonal entries of the response matrix as our data, or the diagonal entries of the rank one matrix formed from any eigenvector (computed or measured directly) of the time-reversal matrix as our data. We assume at first that there is a single target present. Then, for the eigenvectors, the diagonal entries will be real and positive numbers, but they will all contain a constant normalization factor associated with the norm of the eigenvector. For the response matrix, the diagonal entries will be complex and contain a constant factor associated with the scattering strength of the target. We eliminate the unit magnitude phase factor in response matrix diagonals by taking the magnitudes of these entries. Then, we see that in all the cases considered these diagonal data have (for homogeneous 3D media) the form

$$f_n = \frac{\gamma_i}{(4\pi)^2 |\mathbf{r}_n - \mathbf{r}_i|^2} \quad \text{for } n = 1, \dots, N, \quad (29)$$

where γ_i is the magnitude of the scattering strength for the response matrix data, or the normalizing constant for the eigenvector data. The location of the target is \mathbf{r}_i and the location of the n th element of the acoustic array (having a total of N elements) is \mathbf{r}_n .

Using these data, we want to construct a figure of merit that will identify the target location by producing either a noticeably high or a noticeably low value for any point in the imaging region to scanned. To accomplish this, we form the numbers

$$\phi_n(\mathbf{r}) = f_n (4\pi)^2 |\mathbf{r} - \mathbf{r}_n|^2 \quad (30)$$

where \mathbf{r} is the location of any point in the imaging region. Then, we see that when $\mathbf{r} = \mathbf{r}_i$ is located at or very near to the target, then

$$\phi_n(\mathbf{r}_i) = \gamma_i \quad (31)$$

for all N functions ϕ_n . So we want to construct a figure of merit that gives special significance to functions that are positive and constant. But it is precisely this feature that distinguishes the maximum-entropy approach to imaging. If we define an entropy functional H such that

$$H(p_1, \dots, p_N) = -k \sum_{n=1}^N p_n \ln p_n, \quad (32)$$

with the constraint that the probabilities $p_n \geq 0$ and $\sum p_n = 1$, then we construct a maximum principle based on the cost or objective functional

$$J(p_1, \dots, p_N, \lambda) = H(p_1, \dots, p_N) + \lambda \left(\sum p_n - 1 \right), \quad (33)$$

where λ is the Lagrange multiplier for the constraint. The maximum occurs when the constraint is satisfied, and when

$$p_n = e^{(\lambda - k)/k} \quad \text{for all } n = 1, \dots, N. \quad (34)$$

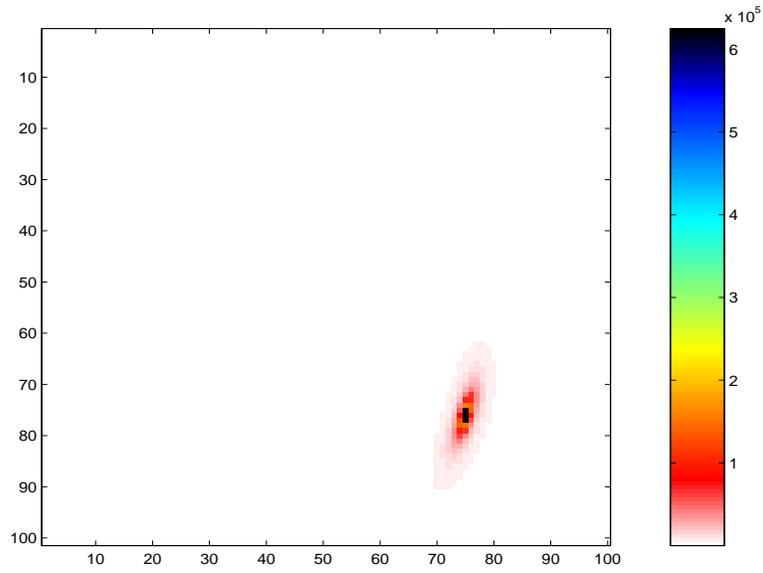


Figure 1: Time-reversal imaging for a single target using (38). jim1-scan1ceps [NR]

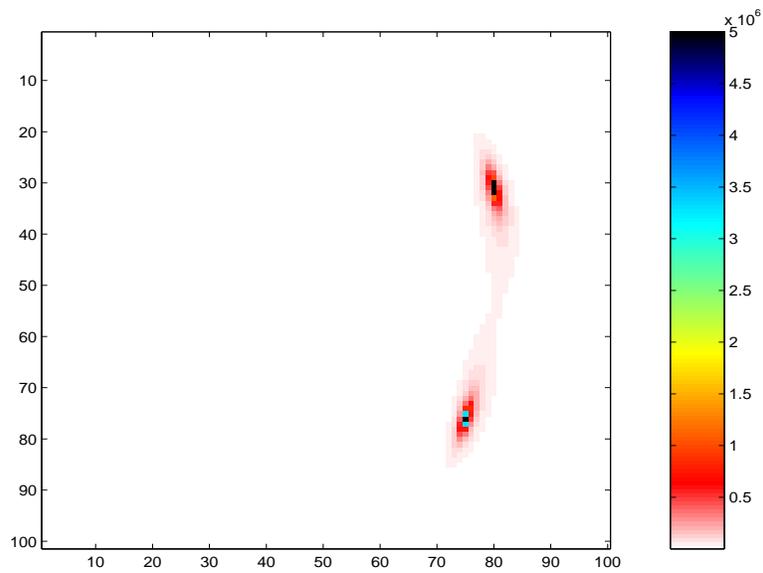


Figure 2: Time-reversal imaging for two targets using (38). jim1-scan34 [NR]

Thus, the maximum entropy occurs when all N states are equally probable.

We can turn this useful fact into an imaging principle by making a small modification in the foregoing derivation. If we define

$$p_n \equiv \phi_n/c, \quad (35)$$

where $c \equiv \sum_n \phi_n$, we see that the maximum-entropy functional can be used as a means of identifying spatial locations at which the various ϕ_n values converge to a constant. The constant k is not important for this application and can be taken as unity. The value of the Lagrange multiplier at the maximum can be determined using (31) and (34) to be

$$\lambda = 1 + \ln(\gamma_i/c) = 1 - \ln(N), \quad (36)$$

since $c = N\gamma_i$ at the target location. Near the maximum of J , we can approximate it in either of two ways:

$$J \simeq -\frac{1}{N} \sum_n \left(\frac{\phi_n}{\gamma_i} - 1 \right)^2 = 1 - \frac{1}{N} \sum_n \left(\frac{\phi_n}{\gamma_i} \right)^2$$

or

$$J \simeq -\frac{1}{N} \sum_n \left[\ln \left(\frac{\phi_n}{\gamma_i} \right) \right]^2. \quad (37)$$

For our present purposes, the second of these forms has proven to be somewhat preferable over the other. The normalizing constant N in this expression has no effect on the result, and whether we look for the minimum or maximum of our function is an arbitrary choice, so we can choose instead to study

$$\tilde{J} \equiv \sum_n \left[\ln \left(\frac{\phi_n}{\gamma_i} \right) \right]^2. \quad (38)$$

which still requires that we know the scattering strength or norming constant γ_i , which we may or may not know. Figures 1 and 2 show the results obtained for one and two targets using (38).

If we know γ_i , then we can image the target using (38) directly. If we do not know γ_i , then we need an estimate of it. One convenient way of obtaining an estimate is by picking any one of the ϕ_n values as the estimate. Clearly, this choice gives a good approximation to the right result at the target, but it will also cause some smearing of the image. The imaging algorithm in this case is then based upon

$$\tilde{J}_q \equiv \sum_n \left[\ln \left(\frac{\phi_n}{\phi_q} \right) \right]^2, \quad (39)$$

where q is any one of the values $n = 1, \dots, N$. This approach works and gives the results shown in Figures 3 and 4 (which should then be compared and contrasted with the results in Figures 1 and 2). In these Figures, we chose q to be the transducer coordinate of the one

that measured the largest amplitude of all the transducers. We see that the results are a little peculiar in the sense that the region of disturbed values near the target location has a teardrop shape, and the center of the teardrop also has some curvature directed away from the center of the array. This observation suggests that it might be preferable not to make any particular choice of q , but instead to consider them all equally. We can do so by symmetrizing the result as much as possible with data available, and also possibly sharpen the image. This criterion results in the imaging objective functional

$$\hat{J} \equiv \sum_{n,q} \left[\ln \left(\frac{\phi_n}{\phi_q} \right) \right]^2 \quad (40)$$

The result of using this criterion is shown in Figure 5, which should be compared directly to Figure 4.

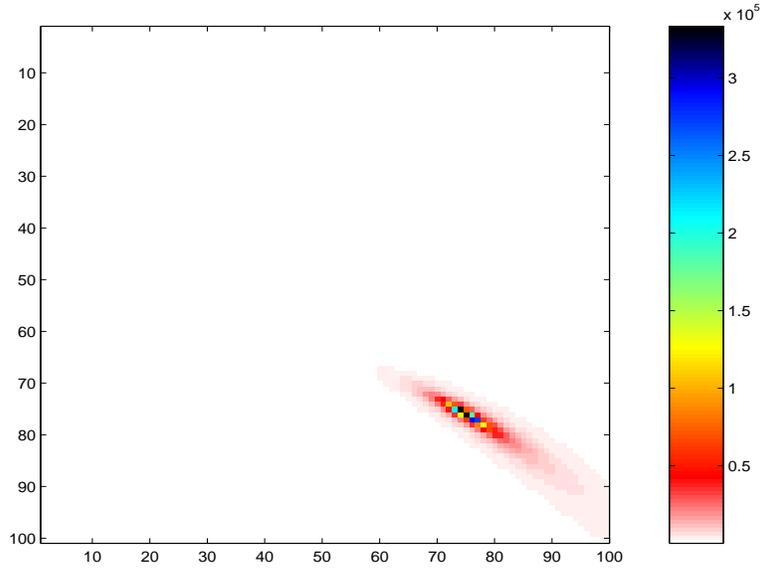


Figure 3: Response matrix imaging for a single target using (39). `jim1-scan2ceps` [NR]

To understand a little better what this symmetrized maximum-entropy imaging scheme is doing to map the data into an image, we will expand (40) so that

$$\begin{aligned} \hat{J} &= \sum_{n,q} [\ln \phi_n - \ln \phi_q]^2 \\ &= 2 \left[\sum_{n,q} (\ln \phi_n)^2 - \sum_n \ln \phi_n \sum_q \ln \phi_q \right] \\ &= 2 \left[N \sum_n (\ln \phi_n)^2 - \left(\sum_n \ln \phi_n \right)^2 \right] \end{aligned} \quad (41)$$

By defining an averaging operator over the functional values at the locations of the N transducers such that $\langle \cdot \rangle \equiv \frac{1}{N} \sum \cdot$, we see that (41) is of the form

$$\hat{J} = 2N^2 [\langle |\ln \phi|^2 \rangle - \langle \ln \phi \rangle^2] \quad (42)$$

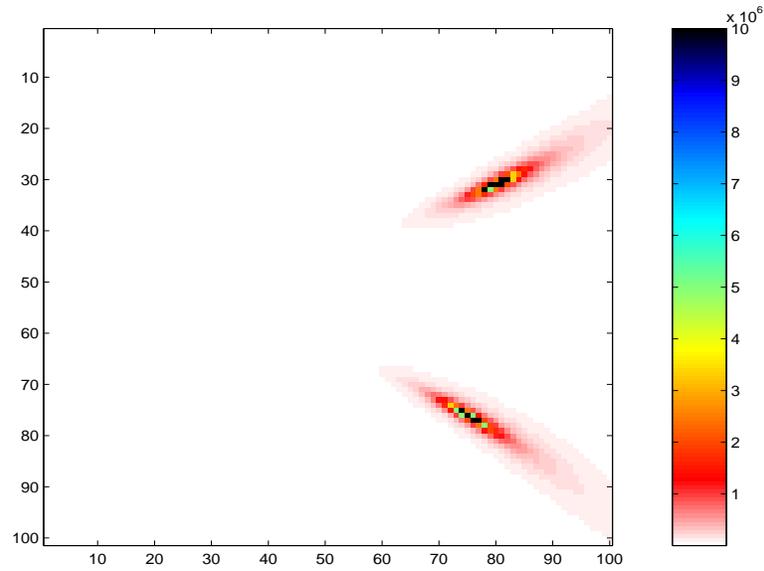


Figure 4: Response matrix imaging for two targets using (39). jim1-scan35 [NR]

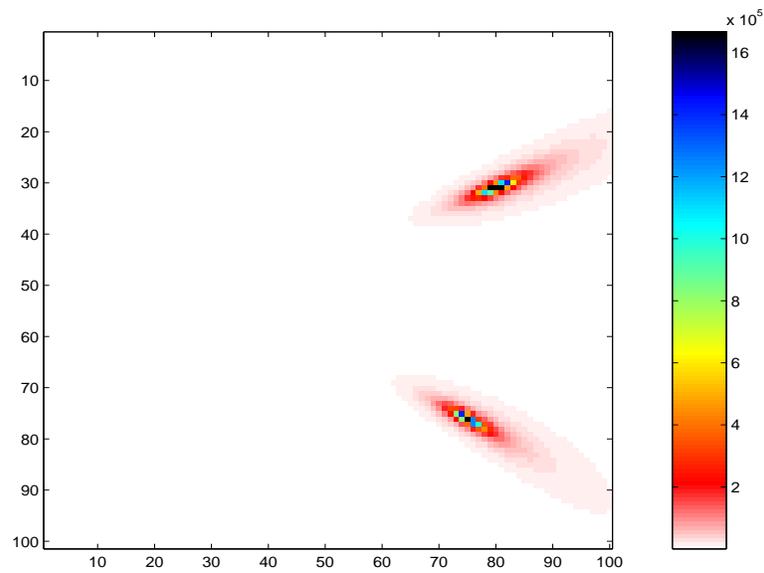


Figure 5: Response matrix imaging for two targets using (40). jim1-scan45 [NR]

and therefore shows that \hat{J} is a measure of the fluctuations in $|\ln\phi|$ at each location in the region being mapped. At the target location, the fluctuations vanish identically, since they become equal to the constant $\ln\gamma_i$. Equation (42) is very useful for two reasons: (1) it shows how the modified maximum entropy imaging criterion is related to fluctuations in ϕ , and (2) it points out that the form \hat{J} could have been postulated as our imaging criterion in the first place, independent of the derivation provided here, since it uses exactly the same features of the data to distinguish the location of the target.

CONCLUSIONS

After describing a well-known method in radio astronomy called maximum-entropy imaging, I introduced a new method that also uses the maximum-entropy concept in a new way. The main distinction between these two approaches is that the well-known approach uses the entropy functional as a means of regularizing the data inversion in order to produce well-resolved images of point objects – stars in the case of radio astronomy. The new method uses the entropy functional not as a smoother or regularizer, but rather as the imaging criterion. When the entropy functional is maximum, or equivalently when our modified functional vanishes at some point in space, that is the location of a target. By plotting the inverse of this modified functional, we arrive at a method that has much in common with the MUSIC algorithm for imaging, but the new method uses different data. MUSIC requires the computation of the SVD of the response matrix. The new method can make use of this information if available, but does not require it. Furthermore, the new method can make use of just the diagonal elements of the response matrix (as in synthetic aperture imaging), or it can be used with time domain amplitude data for the main arrivals.

Future work on this approach will explore how robust the method is when the propagating medium is itself random, in addition to the presence of the isolated scatterers/targets we want to localize.

ACKNOWLEDGMENTS

I benefited greatly from discussions of these and related imaging methods with Liliana Borcea, George Papanicolaou, Chrysoula Tsogka, and Michael Vogelius. I thank Brad Artman, Morgan Brown, and David Chambers for helpful comments on the manuscript.

REFERENCES

- Blomgren, P., Papanicolaou, G. C., and Zhao, H., 2001, Super-resolution in time-reversal acoustics: *J. Acoust. Soc. Am.*, to appear.
- Brühl, M., Hanke, M., and Vogelius, M. S., 2001, A direct impedance tomography algorithm for locating small inhomogeneities: preprint.

- Chambers, D. H., and Gautesen, A. K., 2001, Time reversal for a single spherical scatterer: *J. Acoust. Soc. Am.* **109**, 2616–2624.
- Cherkaeva, E., and Tripp, A. C., 1996a, Optimal survey design using focused resistivity arrays: *IEEE Trans. Geosci. Remote Sensing* **34**, 358–366.
- Cherkaeva, E., and Tripp, A. C., 1996b, Inverse conductivity problem for inexact measurements: *Inverse Problems* **12**, 869–883.
- Colton, D., and Kirsch, A., 1996, A simple method for solving inverse scattering problems in the resonance region: *Inv. Prob.* **12**, 383–393.
- Devaney, A. J., 1999, Super-resolution processing of multi-static data using time reversal and MUSIC: preprint.
- Fink, M., 1997, Time reversal acoustics: *Physics Today* **50**, 34–40 (March, 1997).
- Fink, M., 1999, Time-reversed acoustics: *Scientific American* **281**, 91–97 (November, 1999).
- Gisser, D. G., Isaacson, D., and Newell, J. C., 1990, Electric current computed tomography and eigenvalues: *SIAM J. Appl. Math.* **50**, 1623–1634.
- Groetsch, C. W., 1984, *Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, Massachusetts.
- Gull, S. F., and Daniell, G. J., 1978, Image reconstruction from incomplete and noisy data: *Nature* **272**, 686–690.
- Gull, S. F., and Skilling, J., 1984, Maximum entropy imaging method in image processing: *IEE Proc. – Part F. Commun., Radar, Signal Proc.* **131**, 646–659.
- Haber, E., Ascher, U. M., and Oldenburg, D., 2000, On optimization techniques for solving nonlinear inverse problems: *Inverse Problems* **16**, 1263–1280.
- Hanke, M., 1997, A regularizing Levenberg-Marquardt scheme with applications to Inverse groundwater filtration problems: *Inverse Problems* **13**, 79–95.
- Högbom, J. A., 1974, Aperture synthesis with a non-regular distribution of interferometer baselines: *Astron. Astrophys. Suppl.* **15**, 417–426.
- Isaacson, D., 1986, Distinguishability of conductivities by electric current computed tomography: *IEEE Trans. Med. Imaging* **MI-5**, 91–95.
- Jaynes, E. T., 1957, Information theory and statistical mechanics: *Phys. Rev.* **106**, 620–630.
- Kirsch, A., 1998, Characterization of the shape of a scattering obstacle using the spectral data of the far field operator: *Inv. Prob.* **14**, 1489–1512.
- Marple, S. L., 1987, *Digital Spectral Analysis*, Prentice-Hall, Englewood Cliffs, NJ, pp. 373–378.

- Mast, T. D., Nachman, A. I., and Waag, R. C., 1997, Focusing and imaging using eigenfunctions of the scattering operator: *J. Acoust. Soc. Am.* **102**, 715–725.
- Morozov, V. A., 1967, On the solution of functional equations by the method of regularization: *Sov. Math. Doklady* **7**, 414–417.
- Morozov, V. A., 1984, *Methods for Solving Incorrectly Posed Problems*, Springer, New York.
- Prada, C., and Fink, M., 1994, Eigenmodes of the time reversal operator: A solution to selective focusing in multiple-target media: *Wave Motion* **20**, 151–163.
- Prada, C., Thomas, J.-L., and Fink, M., 1995, The iterative time reversal process: Analysis of the convergence: *J. Acoust. Soc. Am.* **97**, 62–71.
- Prada, C., Wu, F., and Fink, M., 1991, The iterative time reversal mirror: A solution to self-focusing in the pulse echo mode: *J. Acoust. Soc. Am.* **90**, 1119–1129.
- Schmidt, R. O., 1979, Multiple emitter location and signal parameter estimation: *Proc. RADC Spectrum Estimation Workshop*, Rome Air Development Center, Rome, NY, RADC-TR-79-63, October, 1979, pp. 243–258 [reprinted in *IEEE Trans. Antennas Prop.* **AP-34**, 276–280 (1986)].
- Stoica, P., and Moses, R. L., 1997, *Introduction to Spectral Analysis*, Prentice-Hall, Englewood Cliffs, NJ.
- Stoica, P., and Nehoral, A., 1990, MUSIC, maximum likelihood, and Cramér-Rao bound: Further results and comparisons: *IEEE Trans. Acoustics, Speech, and Signal Proc.* **38**, 2140–2150.
- Tikhonov, A. N., and Arsenin, V. Ya., 1979, *Methods for Solving Ill-Posed Problems* Wiley, New York.
- Tsogka, C., and Papanicolaou, G. C., 2001, Time reversal through a solid-liquid interface and super-resolution: *J. Acoust. Soc. Am.* , submitted for publication.
- Xu, W., and Kaveh, M., 1996, Comparative study of the biases of MUSIC-like estimators: *Signal Processing* **50**, 39–55.

Stolt residual migration for converted waves

Daniel Rosales, Paul Sava, and Biondo Biondi¹

ABSTRACT

PS velocity analysis is a new frontier in converted-waves seismic imaging. To obtain one consistent image, it is necessary to estimate correctly both the *P*-waves velocity model and the *S*-waves velocity model. Stolt residual migration is a useful technique for image update and velocity analysis. This paper extends Stolt prestack residual migration in order to handle two different velocity fields. The operator that we introduce is promising for *PS* velocity analysis. We present a theoretical discussion of our new operator, and discuss its ability to focus *PS* images.

INTRODUCTION

Stolt residual prestack migration is a useful tool to improve the quality of the image and to perform migration velocity analysis (Stolt, 1996; Sava, 1999, 2000). To update a *PS* image, we extend Stolt prestack residual migration for two different velocity fields.

The extension of Stolt residual migration for converted waves can be done in two ways. An approximate method uses two different residual velocities in each of the two square roots of the double square root equation. The exact method calculates an appropriate transformation kernel that is able to handle both velocity fields. We will discuss both ways in the theory part. Moreover, we present synthetic examples, with both constant velocity and depth variant velocity models. For constant velocity, we show that we are able to focus the image after a migration with wrong velocities. For depth-variant velocity, we show that we are able to improve the image.

THEORY

Stolt (1996) first introduced prestack residual migration. Sava (1999) reformulated Stolt residual migration in order to handle prestack depth images. This section presents the extension of Sava (1999) for two different wavefields, therefore, two different velocities. We present this extension for converted waves data, where the *P* to *S* conversion occurs at the reflector. Although the formulation involves *P*-velocities and *S*-velocities, its application is not limited to converted waves only. Rosales and Biondi (2001) present a possible application for imaging under salt edges.

¹email: daniel@sep.stanford.edu,paul@sep.stanford.edu,biondo@sep.stanford.edu

Residual prestack Stolt migration operates in the Fourier domain. Considering the representation of the input data in shot-geophone coordinates, the mapping from the data space to the model space takes the form

$$k_z = \frac{1}{2} \left(\sqrt{\frac{\omega^2}{v_p^2} - k_s^2} + \sqrt{\frac{\omega^2}{v_s^2} - k_g^2} \right), \quad (1)$$

where k_s , k_g , v_p , and v_s stand for, respectively, the source and geophone wavenumbers, and the P and S velocities.

In residual prestack Stolt migration for converted waves, we attempt to simultaneously correct the effects of migrating with two inaccurate velocity fields.

Supposing that the initial migration was done with the velocities v_{0p} and v_{0s} , and that the correct velocities are v_{mp} and v_{ms} , we can then write

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

Solving for ω^2 in the first equation of (2) and substituting it in the second equation of (2), we obtain the expression for prestack Stolt depth residual migration for converted waves [equations (3) and/or (4)] (see Appendix A, for details in derivation)

$$k_{z_m} = \frac{1}{2} \sqrt{\rho_p^2 \overline{\kappa_0^2} - k_s^2} + \frac{1}{2} \sqrt{\rho_s^2 \gamma_0^2 \overline{\kappa_0^2} - k_g^2}, \quad (3)$$

$$k_{z_m} = \frac{1}{2} \sqrt{\rho_p^2 \overline{\kappa_0^2} - k_s^2} + \frac{1}{2} \sqrt{\rho_p^2 \gamma_m^2 \overline{\kappa_0^2} - k_g^2}, \quad (4)$$

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

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

and $\rho_p = \frac{v_{0p}}{v_{mp}}$, $\rho_s = \frac{v_{0s}}{v_{ms}}$, and $\gamma = \frac{v_p}{v_s}$.

In equation (2) it appears that Stolt residual migration for converted waves depends on four parameters: v_{0p} , v_{0s} , v_{mp} , v_{ms} . These four degrees of freedom can be reduced to three (ρ_p , ρ_s and γ), as seen in equations (3) and (4) and demonstrated in Appendix A. This is important, because a three parameters search for updating converted waves images is simpler than a four parameters search. However, it would be useful to further reduce the number of parameter to two. Assuming that the v_p/v_s ratio is the same after and before the residual migration process, it is possible to simplify equations (3) and/or (4) into a two parameter equation:

$$k_{z_m} = \frac{1}{2} \sqrt{\rho_p^2 \overline{\kappa_0^2} - k_s^2} + \frac{1}{2} \sqrt{\rho_p^2 \gamma^2 \overline{\kappa_0^2} - k_g^2}, \quad (5)$$

where the transformation kernel takes the form

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

If we just specify two different ratios in both square roots of Sava's (1999) formulation we have

$$k_{z_m} = \frac{1}{2}\sqrt{\rho_p^2 \kappa_0^2 - k_s^2} + \frac{1}{2}\sqrt{\rho_s^2 \kappa_0^2 - k_g^2}, \quad (6)$$

where the transformation kernel has the same form as the one of PP waves:

$$\kappa_0^2 = \frac{[4k_{z_0}^2 + (k_g - k_s)^2][4k_{z_0}^2 + (k_g + k_s)^2]}{16k_{z_0}^2}.$$

Equation (6) shows another way of doing prestack residual migration for converted waves. Although equations (5) and (6) may look similar because they depend on only two parameters, the transformation kernels (κ_0 and $\bar{\kappa}_0$) are different. Equation (6) has a similar transformation kernel as the conventional PP prestack residual migration, while equation (5) presents a kernel deduced for the case of converted waves.

It is important to note that all three equations (3), (5), and (6) reduce to the same expressions in the limit when of v_s tends to v_p , or γ tends to 1. All of them reduce to the simple case of prestack residual migration for conventional PP data. Appendix B demonstrates this result.

NUMERICAL EXAMPLES

Impulse Responses

Figure 1 shows the impulse response of the three residual prestack migration operators [equations (3), (5), (6)].

Figure 1a presents the impulse response for equation (3), with $\rho_p = 1.2$, $\rho_s = 1.2$ and $\gamma_0 = 2$. Figure 1b presents the impulse response for equation (5), with $\rho_p = 1.2$ and $\gamma = 2$. As expected from the theory discussed in the previous section, Figure 1a and Figure 1b are identical because $\gamma = \gamma_0$.

Figure 1c presents the impulse response for equation (6), with $\rho_p = 1.2$, $\rho_s = 1.2$. It is possible to observe the difference with respect to figures 1a and 1b. The difference is due to the approximation in the transformation kernel.

Figure 2 demonstrates the differences between equations (3) and (5). Figure 2a shows the impulse response for equation (3) with, $\rho_p = 0.9$, $\rho_s = 1.4$ and $\gamma_0 = 2$. Figure 2b shows the impulse response for equation (5) with, $\rho_p = 0.9$ and $\gamma = 2$. It is easy to observe the difference between the impulse responses due to the approximation in the transformation kernel.

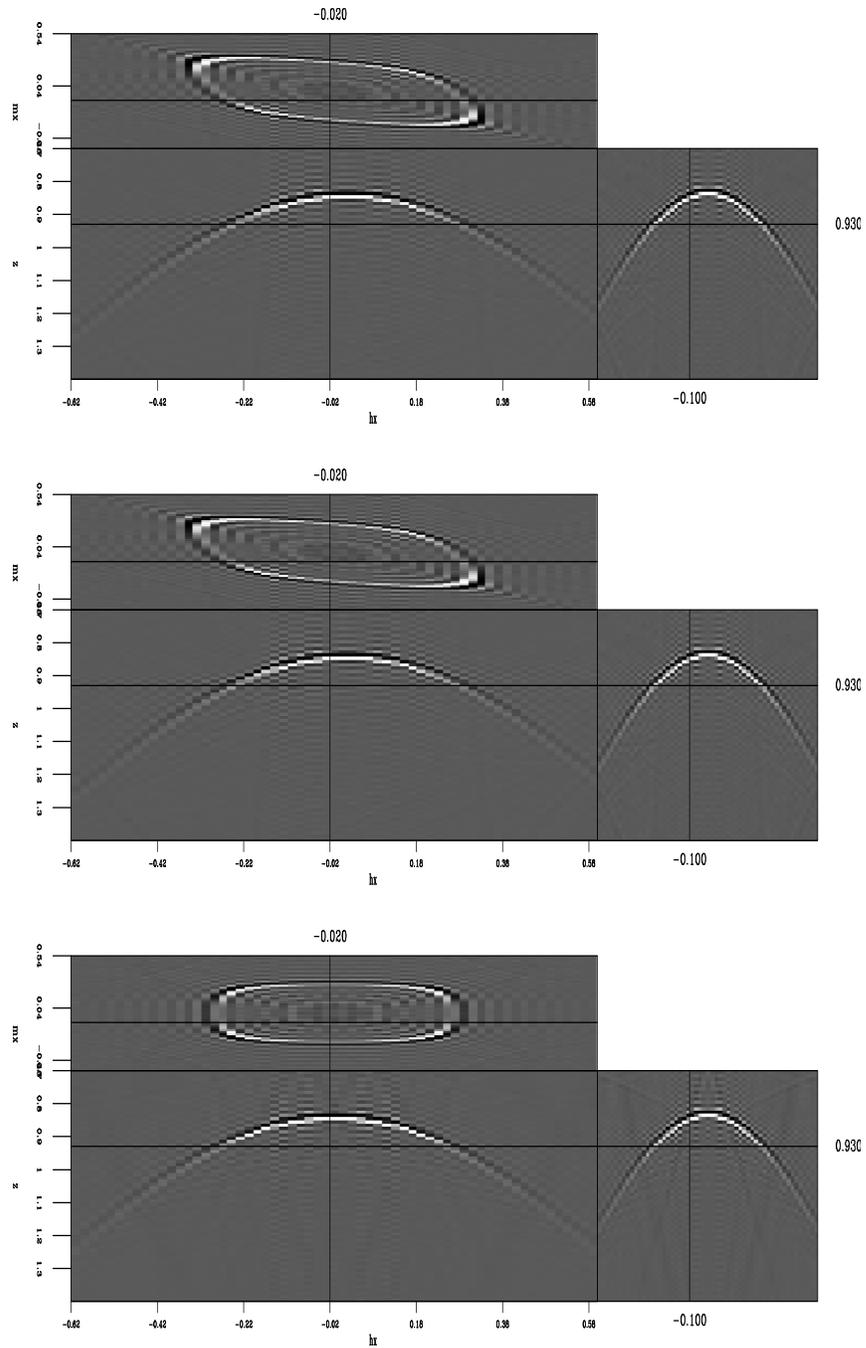


Figure 1: Impulse response for the residual prestack migration operators. From top to bottom, a) equation (3) for $\rho_p = 1.2$, $\rho_s = 1.2$, $\gamma_0 = 2$; b) equation (5) for $\rho_p = 1.2$, $\gamma = 2$; c) equation (6) for $\rho_p = 1.2$, $\rho_s = 1.2$. `daniel1-imp` [ER,M]

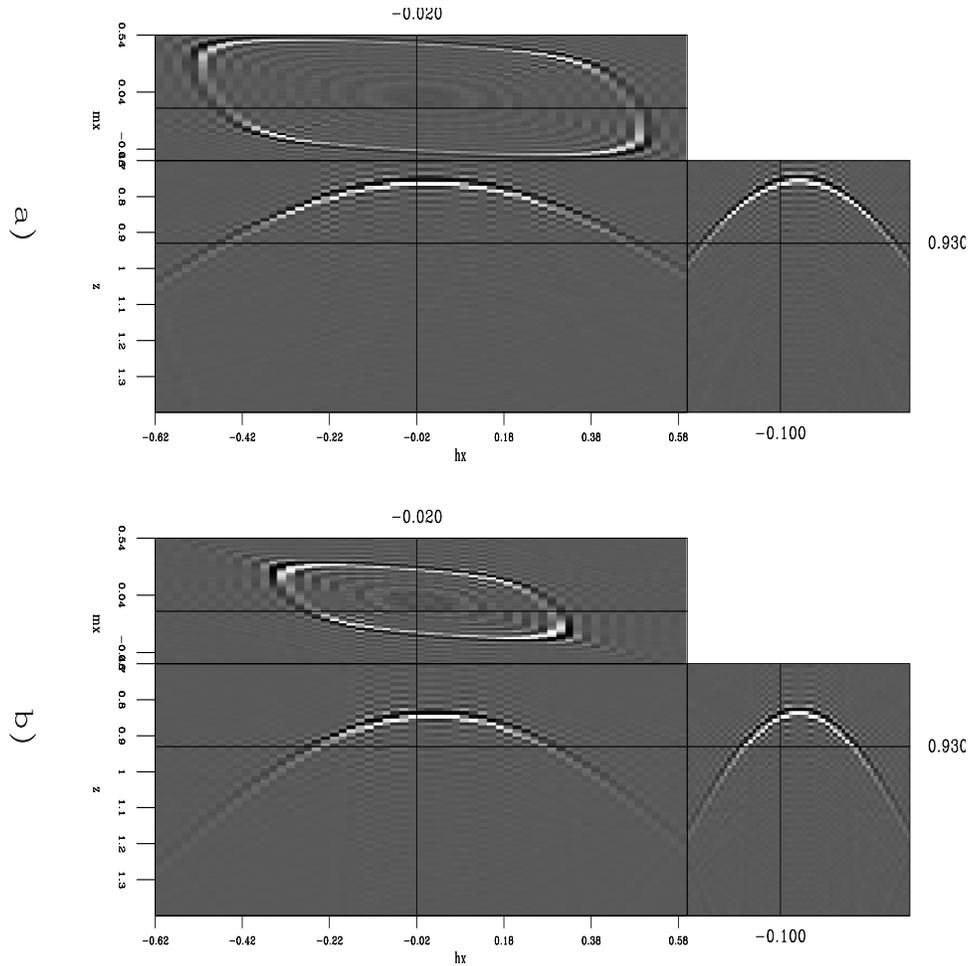


Figure 2: Impulse response for the residual prestack migration operators. From top to bottom, a) equation (3) for $\rho_p = 1.2$, $\rho_s = 1.4$, $\gamma_0 = 2$; b) equation (5) for $\rho_p = 1.2$, $\gamma = 2$. `daniel1-imp2` [ER,M]

Residual migration with constant velocity

In order to test the accuracy of our operator we created a synthetic data set assuming a constant velocity model of $v_p = 3000$ m/s and $v_s = 1500$ m/s, and four reflectors having dips of 15° , 0° , -30° , -45° , respectively.

We first migrated the data with the wrong velocity, using shot-profile migration for converted waves (Rosales and Rickett, 2001), and extract angle-domain common-image gathers for converted waves (Rickett and Sava, 2001). We then applied Stolt residual prestack migration, using the methods described in the previous section.

We show the results of performing migration followed by residual prestack migration, when the initial migration was performed with the correct P -velocity, and with the S -velocity 15% too high. Figure 3 shows a comparison for the zero offset section, equivalent to the

stack final migrated section, of the residual migration result for equations (3), (5) and (6) and the migration with the correct velocity model. It is possible to observe that the result with the correct expression [equation(3)] produces an image similar to the image produced by the migration with the correct velocity. On the contrary, when we use the other two expressions we obtain poorly focused results.

Notice that because of the change in polarity between the negative and positive reflection angles, these zero-offset sections show degraded images. Rosales and Rickett (2001) discuss how to correct for the change in polarity. Therefore, we analyze the common image gathers. Figure 4 shows five common image gather all taken at the same surface location. From left to right: the good migration result, correction with the correct expression [equation (3)], correction with equation (5), correction with equation (6) and the bad migration result. It is possible to observe that the events after residual migration with the correct expression [equation (3)] are flat. Corrections with equations (5) and (6) yield to over and under corrected gathers, respectively.

Residual migration with vertical velocity gradient

Since Stolt residual prestack migration is based on the assumption of constant velocity, we will evaluate the accuracy for velocity models having depth dependence.

We create our second synthetic model using the same reflector geometry as the previous section and using a realistic depth velocity model of $v_{0p} = 1700$ m/s with a gradient of 0.15 s⁻¹ and $v_{0s} = 300$ m/s with a gradient of 0.35 s⁻¹,

As before, we perform the initial shot profile migration with a 15% positive perturbation in the S -velocity model. We perform residual migration with the RMS γ_0 value at approximately 1.5 km, and $\rho_p = 1$ and $\rho_s = 1.15$. Residual prestack migration could not recover the image correctly at all depths. However, we got a good approximation of it. Figures 5 and 6 show our results.

CONCLUSIONS

We introduced the extension of Stolt prestack residual migration for converted waves. Our new operator involves the selection of three parameters in order to update the image.

To help in the memory and disk space necessary for the implementation of our operator, we also derived approximations that reduce the number of free parameters to a two. The most appropriate way of reducing the number of parameters is by freezing one of them. Our experience suggests freezing γ_0 at the RMS value of the ratio between the P and the S migration velocities.

In constant velocity, we proved that we can recover the image obtained with an initial migration that uses an inaccurate velocity model. Therefore, we can update a migration with constant two-velocities model using our new operator. We can also update an image obtained

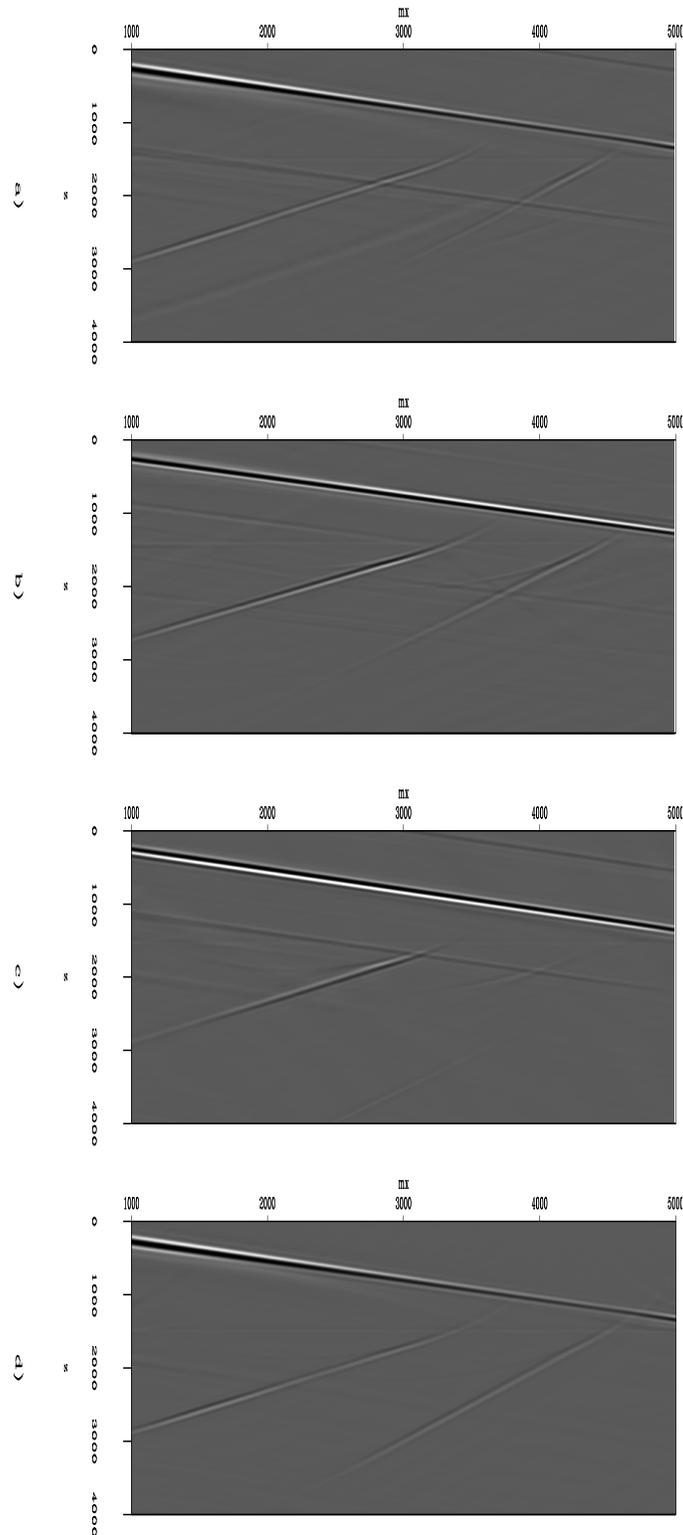


Figure 3: Zero offset section, equivalent to final migrated section, using the residual migration operator for: a) equation (3) for $\rho_p = 1$ $\rho_s = 1.15$, $\gamma_0 = 1.74$; b) equation (5) for $\rho_p = \rho_s = 1.15$ and $\gamma = 1.74$; c) equation (6) for $\rho_p = 1$ $\rho_s = 1.15$, d) migration with the correct velocity model. [daniel1-cv1](#) [CR,M]

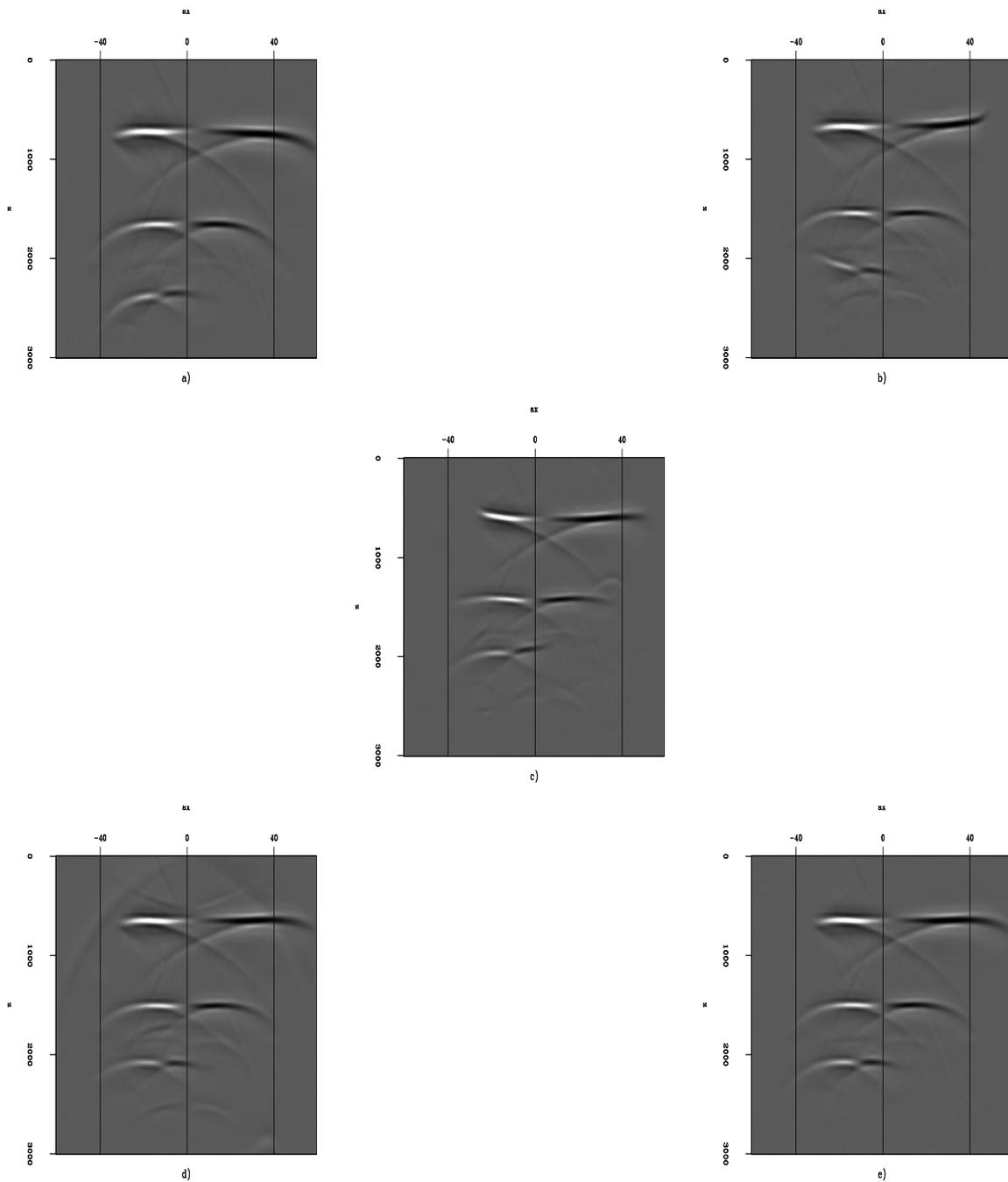


Figure 4: Angle domain common image gathers for, top to bottom and left to right: a) migration with 15% error in the S -velocity; b) residual migration with equation (6); c) residual migration with equation (5); d) residual migration with equation (3); e) migration with the correct velocity. [daniel1-cv2](#) [CR,M]

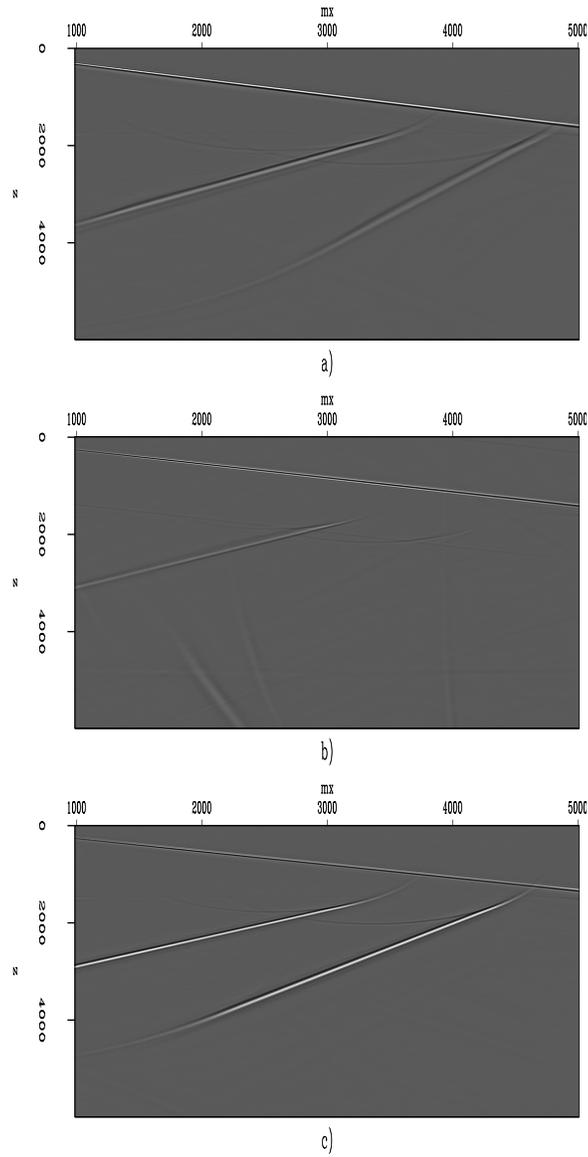


Figure 5: Zero offset section comparison, final migration stack, for: a) migration with 15% error in the S -velocity model; b) residual migration with equation (6) for $\rho_p = 1$, $\rho_s = 1.15$ and $\gamma_0 = 1.92$; c) migration with the correct velocity model. [daniel1-vofz](#) [CR,M]

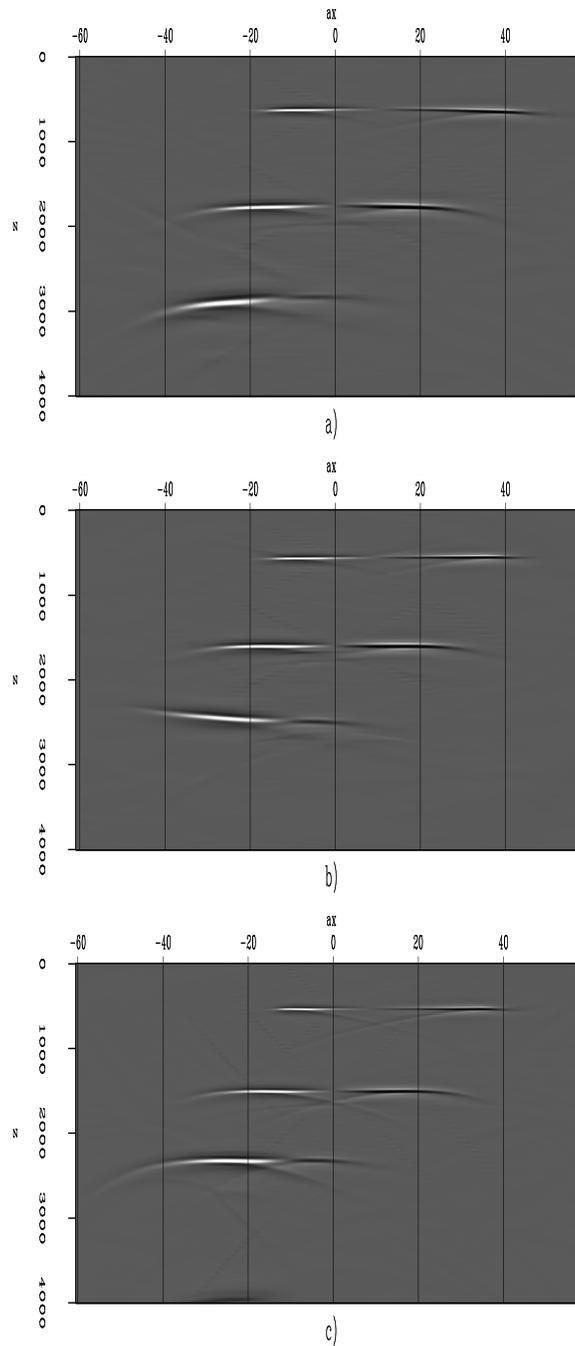


Figure 6: Zero offset section comparison, final migration stack, for: a) migration with 15% error in the S -velocity model; b) residual migration with equation (6) for $\rho_p = 1$, $\rho_s = 1.15$ and $\gamma_0 = 1.92$; c) migration with the correct velocity model. [daniel1-vofz2](#) [CR,M]

with a depth variant velocity. However, the refocusing performed by residual migration is only approximate.

The advantages of having an operator to update converted waves images allow us to extrapolate our ability to handle multiple mode data. We hope that it will lead to more accurate methods for performing velocity analysis for converted waves.

APPENDIX A

Solving the first equation of:

$$\begin{cases} k_{z_0} = \frac{1}{2} \left(\sqrt{\omega^2 s_{p_0}^2 - k_s^2} + \sqrt{\omega^2 s_{s_0}^2 - k_g^2} \right) \\ k_{z_m} = \frac{1}{2} \left(\sqrt{\omega^2 s_{p_m}^2 - k_s^2} + \sqrt{\omega^2 s_{s_m}^2 - k_g^2} \right), \end{cases} \quad (\text{A-1})$$

for ω^2 we have:

$$2k_z = \sqrt{\omega^2 s_{p_0}^2 - k_s^2} + \sqrt{\omega^2 s_{s_0}^2 - k_g^2},$$

and

$$\left(2k_z - \sqrt{\omega^2 s_{p_0}^2 - k_s^2} \right)^2 = \left(\sqrt{\omega^2 s_{s_0}^2 - k_g^2} \right)^2.$$

Squaring the previous equation and isolating the remaining square root we obtain:

$$4k_z^2 + \omega^2 s_{p_0}^2 - \omega^2 s_{s_0}^2 + k_g^2 - k_s^2 = 4k_z \sqrt{\omega^2 s_{p_0}^2 - k_s^2}.$$

Squaring the previous equations, grouping common terms, and setting equal to zero, we get:

$$\omega^4 (s_{p_0}^2 - s_{s_0}^2)^2 + \omega^2 [(8k_z^2 + 2k_g^2 - 2k_s^2)(s_{p_0}^2 - s_{s_0}^2) - 16k_z^2 s_{p_0}^2] + (4k_z^2 + k_g^2 - k_s^2)^2 + 16k_z^2 k_s^2 = 0.$$

Solving for ω^2 we obtain

$$\omega^2 = \frac{s_{s_0}^2 (4k_z^2 + k_g^2 - k_s^2) + s_{p_0}^2 (4k_z^2 + k_s^2 - k_g^2) \pm 4k_z^2 \sqrt{s_{p_0}^2 s_{s_0}^2 (4k_z^2 + k_g^2 + k_s^2) - s_{p_0}^4 k_g^2 - s_{s_0}^4 k_s^2}}{(s_{p_0}^2 - s_{s_0}^2)^2}. \quad (\text{A-2})$$

We select the negative sign of the radical as the final solution for ω^2 , as discussed in Appendix B.

Substituting the result of ω^2 in the second equation of relation (A-1), we obtain the relationship for residual prestack migration for converted waves.

In order to demonstrate this fact, we need to simplify the dispersion relation for ω^2 in terms of $\gamma_0 = \frac{v_{p_0}}{v_{s_0}}$, s_{p_0} and s_{s_0} depending on the source or receiver SSR equation.

$$(s_{p_0}^2 - s_{s_0}^2)^2 = s_{p_0}^4 (1 - \gamma_0^2)^2$$

therefore,

$$\omega^2 = \frac{\gamma_0^2(A) + (B) - 4k_z^2 \sqrt{\gamma_0^2(C) - k_g^2 - \gamma_0^4 k_s^2}}{s_{p_0}^2 (1 - \gamma_0^2)^2},$$

since

$$k_{z_m} = \frac{1}{2} \left(\sqrt{\omega^2 s_{p_m}^2 - k_s^2} + \sqrt{\omega^2 s_{s_m}^2 - k_g^2} \right),$$

calling

$$\kappa_0^2 = \frac{\gamma_0^2 A + B - 4k_z^2 \sqrt{\gamma_0^2 C - k_g^2 - \gamma_0^4 k_s^2}}{(1 - \gamma_0^2)^2}, \quad (\text{A-3})$$

we finally note that $\omega^2 = \frac{\kappa_0^2}{s_{p_0}^2}$. We then have equation (3), which is:

$$k_{z_m} = \frac{1}{2} \left(\sqrt{\gamma_0^2 \rho_s^2 \kappa_0^2 - k_g^2} + \sqrt{\rho_p^2 \kappa_0^2 - k_s^2} \right). \quad (\text{A-4})$$

APPENDIX B

We want to evaluate equations (3), (5) and (6) when $v_p = v_s$, or equivalently, when $\gamma_m = \gamma_0 = 1$.

It is possible to see from equation (A-3) that for the particular case of $\gamma_0 = 1$ we have a division by zero. Since we have a division by zero, we need to analyze the equation when we approach to $\gamma_0 \rightarrow 1$. For this purpose, we are going to use L'Hôpital. Therefore, we need to have a zero also in the numerator, which is possible for any value of k_z if, and only if, we choose the negative sign as a solution in equation (A-2).

Referring to equation (A-3) as $\kappa_0^2 = f(\gamma_0)/g(\gamma_0)$, and applying the L'Hôpital, we calculate the derivative with respect to γ_0 to $f(\gamma_0)$ and $g(\gamma_0)$.

We derive

$$\frac{\partial}{\partial \gamma_0} f(\gamma_0) = 2\gamma_0 (4k_z^2 + k_g^2 - k_s^2) - 4k_z^2 \frac{\gamma_0 (8k_z^2 + 2k_g^2 + 2k_s^2 - 4\gamma_0^3 k_s^2)}{\sqrt{\gamma_0^2 (4k_z^2 + k_g^2 + k_s^2) - k_g^2 - \gamma_0^4 k_s^2}}.$$

On the other hand, the derivative of the denominator is:

$$\frac{\partial}{\partial \gamma_0} \left((\gamma_0^2 - 1)^2 \right) = 4(\gamma_0^2 - 1)\gamma_0.$$

Analyzing the limit for $\gamma_0 \rightarrow 1$ of the $f'(\gamma_0)/g'(\gamma_0)$, we still have a $\frac{0}{0}$ relation, which means we must re-apply L'Hôpital,

$$\frac{\partial}{\partial \gamma_0} f'(\gamma_0) = 2(4k_z^2 + k_g^2 - k_s^2) - \frac{k_z \gamma_0 (8k_z^2 + 2k_g^2 + 2k_s^2 - 4\gamma_0^3 k_s^2)^2}{(\gamma_0^2 (4k_z^2 + k_g^2 + k_s^2) k_g^2 - \gamma_0^4 k_s^2)^{\frac{3}{2}}} + \frac{2k_z (8k_z^2 + 2k_g^2 + 2k_s^2 - 12\gamma_0^2 k_s^2)}{(\gamma_0^2 (4k_z^2 + k_g^2 + k_s^2) - k_g^2 - \gamma_0^4 k_s^2)^{\frac{1}{2}}}.$$

On the other hand, the second derivative of the denominator is:

$$\frac{\partial}{\partial \gamma_0} g'(\gamma_0) = 8\gamma_0^2.$$

We finally have

$$\lim_{\gamma_0 \rightarrow 1} \frac{f(\gamma_0)}{g(\gamma_0)} = \lim_{\gamma_0 \rightarrow 1} \frac{\frac{\partial}{\partial \gamma_0} f(\gamma_0)}{\frac{\partial}{\partial \gamma_0} g(\gamma_0)} = \lim_{\gamma_0 \rightarrow 1} \frac{\frac{\partial^2}{\partial \gamma_0^2} f(\gamma_0)}{\frac{\partial^2}{\partial \gamma_0^2} g(\gamma_0)}.$$

Therefore, we have κ_0^2 for $\gamma_0 \rightarrow 1$ reduces to:

$$\left[4k_z^2 + (k_g - k_s)^2\right] \left[4k_z^2 + (k_g + k_s)^2\right]$$

which is the expression for the conventional case of *PP* waves.

REFERENCES

- Rickett, J., and Sava, P., 2001, Offset and angle domain common-image gathers for shot-profile migration: SEP-108, 27–34.
- Rosales, D., and Biondi, B., 2001, On asymmetric Stolt residual migration for imaging under salt edges: SEP-110, 71–74.
- Rosales, D., and Rickett, J., 2001, *ps*-wave polarity reversal in angle domain common-image gathers: SEP-108, 35–44.
- 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.
- Stolt, R. H., 1996, Short note - a prestack residual time migration operator: Geophysics, 61, no. 2, 605–607.

Short Note

Effective AMO implementation in the log-stretch, frequency-wavenumber domain

*Ioan Vlad and Biondo Biondi*¹

INTRODUCTION

Azimuth moveout (AMO), introduced by Biondi et al. (1998), is used as part of the styling goal (in conjunction with a derivative as a roughener) in Biondi and Vlad (2001). This paper describes the implementation of AMO for the above-stated purpose, with a historical background, proof, and discussion of pitfalls and practical steps.

THE AZIMUTH MOVEOUT

AMO is conceived as a cascade of forward and reverse dip moveout (DMO) operators. Thus, the accuracy and speed of the DMO operator used is highly important. Computing the DMO in the frequency domain is accurate and simple, but computationally expensive because the DMO operator is temporally nonstationary. The technique of logarithmic time-stretching, introduced by Bolondi et al. (1982) increases the computational efficiency because the DMO operator is stationary in the log-stretch domain, and Fast Fourier Transforms can be used instead of slow Discrete Fourier Transforms. Gardner (1991), Black et al. (1993) and Zhou et al. (1996) derived equivalent and accurate log-stretch, frequency-wavenumber DMO operators. The implementation of the AMO presented in this paper is based on the derivation and algorithm in Zhou et al. (1996).

THE LOG-STRETCH, FREQUENCY-WAVENUMBER AMO IN 3D

Starting from the parametric DMO relations of Black et al. (1993), Zhou et al. (1996) derives an expression for a DMO applicable on 2D NMO-ed data. In order to extend the expression to 3D, we only have to replace the product kh between the wavenumber and half offset with the dot product of the same quantities, which are vectors in the case of 3D data. In order to perform AMO from the offset \vec{h}_1 to the offset \vec{h}_2 , we need to cascade one forward DMO from

¹email: nick@sep.stanford.edu, biondo@sep.stanford.edu

offset \vec{h}_1 to zero offset with a reverse DMO from zero offset to offset \vec{h}_2 . Thus, applying log-stretch, frequency-wavenumber AMO on a 3D cube of data $P(t, m_x, m_y)|_{\vec{h}_1}$ in order to obtain $P(t, m_x, m_y)|_{\vec{h}_2}$ involves the following sequence of operations:

1. Apply log-stretch along the time axis on the $P(t, m_x, m_y)|_{\vec{h}_1}$ cube, with the formula:

$$\tau = \ln\left(\frac{t}{t_c}\right), \quad (1)$$

where t_c is the minimum cutoff time introduced to avoid taking the logarithm of zero. All samples from times smaller than t_c are simply left untouched, the rest of the procedure will be applied to the cube $P(t > t_c, m_x, m_y)|_{\vec{h}_1}$.

2. 3D forward FFT of the $P(\tau, m_x, m_y)|_{\vec{h}_1}$ cube. The 3D forward Fourier Transform is defined as follows:

$$P(\Omega, k_x, k_y) = \int \int \int P(\tau, m_x, m_y) e^{i(\Omega\tau - k_x m_x - k_y m_y)} d\tau dm_x dm_y \quad (2)$$

It can be seen that the sign of the transform along the τ -axis is opposite to that over the midpoint axes.

3. For each element of the cube, perform the AMO shift:

$$P(\Omega, k_x, k_y)|_{\vec{h}_2} = e^{i(\Phi_1 - \Phi_2)} P(\Omega, k_x, k_y)|_{\vec{h}_1}, \text{ where} \quad (3)$$

$$\Phi_j = \begin{cases} 0, & \text{for } \vec{k} \cdot \vec{h} = 0 \\ \vec{k} \cdot \vec{h}, & \text{for } \Omega = 0 \\ \frac{\Omega}{2} \left\{ \sqrt{1 + \left(\frac{2\vec{k} \cdot \vec{h}}{\Omega}\right)^2} - 1 - \ln \left[\frac{\sqrt{1 + \left(\frac{2\vec{k} \cdot \vec{h}}{\Omega}\right)^2} + 1}{2} \right] \right\} & \text{otherwise} \end{cases}, \quad (4)$$

$$\text{where } \vec{k} \cdot \vec{h} = k_x h_x + k_y h_y \quad (5)$$

and j can take the values 1 or 2. The frequency domain variables must have incorporated in their value a 2π constant (they are defined according to equation (2))

4. Do reverse 3D FFT in order to obtain the $P(\tau, m_x, m_y)|_{\vec{h}_2}$ cube.
5. Do reverse log stretch along the time axis and affix to the top of the cube the slices from times smaller than t_c . The final result is a $P(t, m_x, m_y)|_{\vec{h}_2}$ cube.

Figure 1 shows the impulse response of the above described AMO.

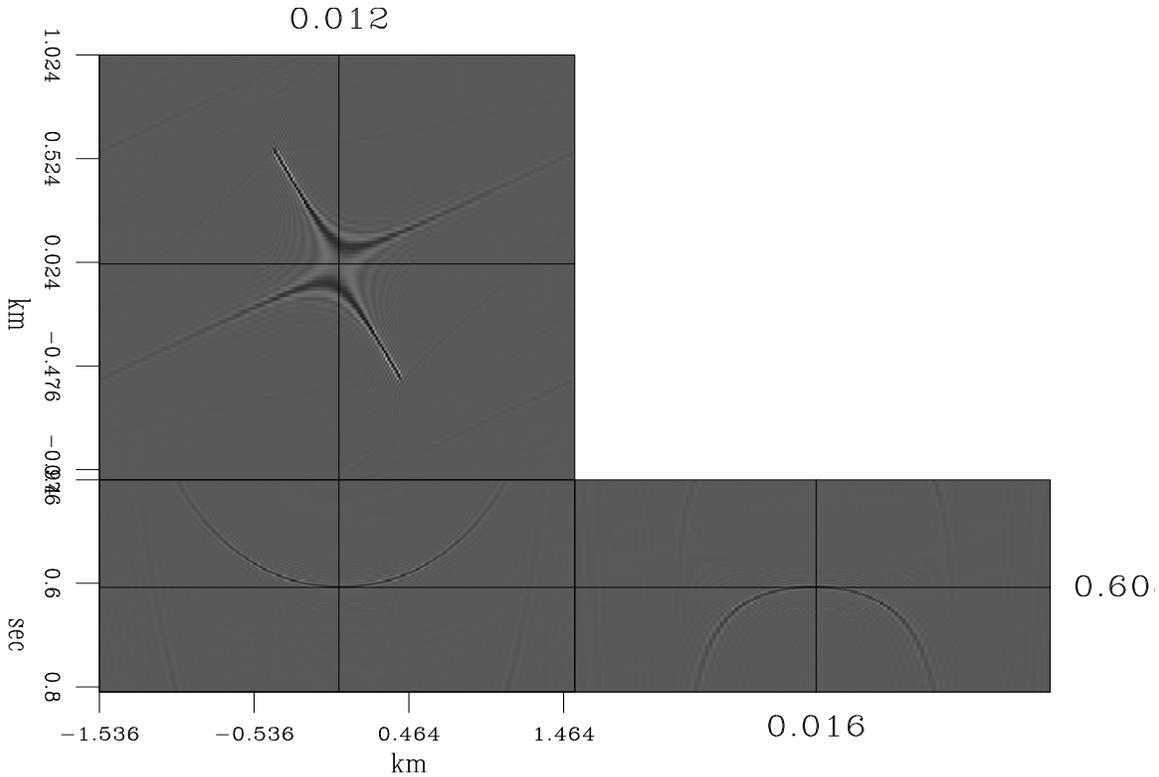


Figure 1: AMO impulse response `nick2-impresp1` [ER]

STRETCHING AND ALIASING

For the purpose of this discussion we define stretching of a single-dimension space as any transformation from one space to another that has the following property: at least an arbitrarily chosen sequence of two consecutive, equal in length, intervals in the input space is transformed into a sequence of two consecutive, *not* equal in length, intervals in the output space. Stretching an x -space to a y -space will be denoted as

$$y = f(x) \quad (6)$$

Two obvious examples of stretching are

$$\begin{aligned} NMO &: y = \sqrt{x^2 + \alpha}, \text{ and} \\ \text{log-stretch} &: y = \log\left(\frac{x}{\alpha}\right), \end{aligned}$$

where α is a positive real number whose value does not matter for the purpose of this discussion. As it can be seen in Fig. 2, if we keep the same sampling rate ($\Delta y = \Delta x$), aliasing can occur when doing the reverse transformation, from x to y . In order to avoid aliasing, we need to compute Δy_{\max} , the largest acceptable sampling rate in the y domain. This can sometimes lead to a larger number of samples in the y domain, and thus to larger computational expense. This can be limited to some extent if the signal in the x -space has been bandpassed, as is often the case with seismic data, with the largest frequency present in the data (f_{\max}) smaller than the

Nyquist frequency given by the sampling rate (f_{Ny}). Thus, we can replace in our calculations Δx with

$$\Delta x_{\max} = \frac{1}{2f_{\max}},$$

which will result in a Δy_{\max} larger than that computed using Δx , the sampling rate in the x space. In order to compute Δy_{\max} , we will consider two points in the x space, as seen in Fig.

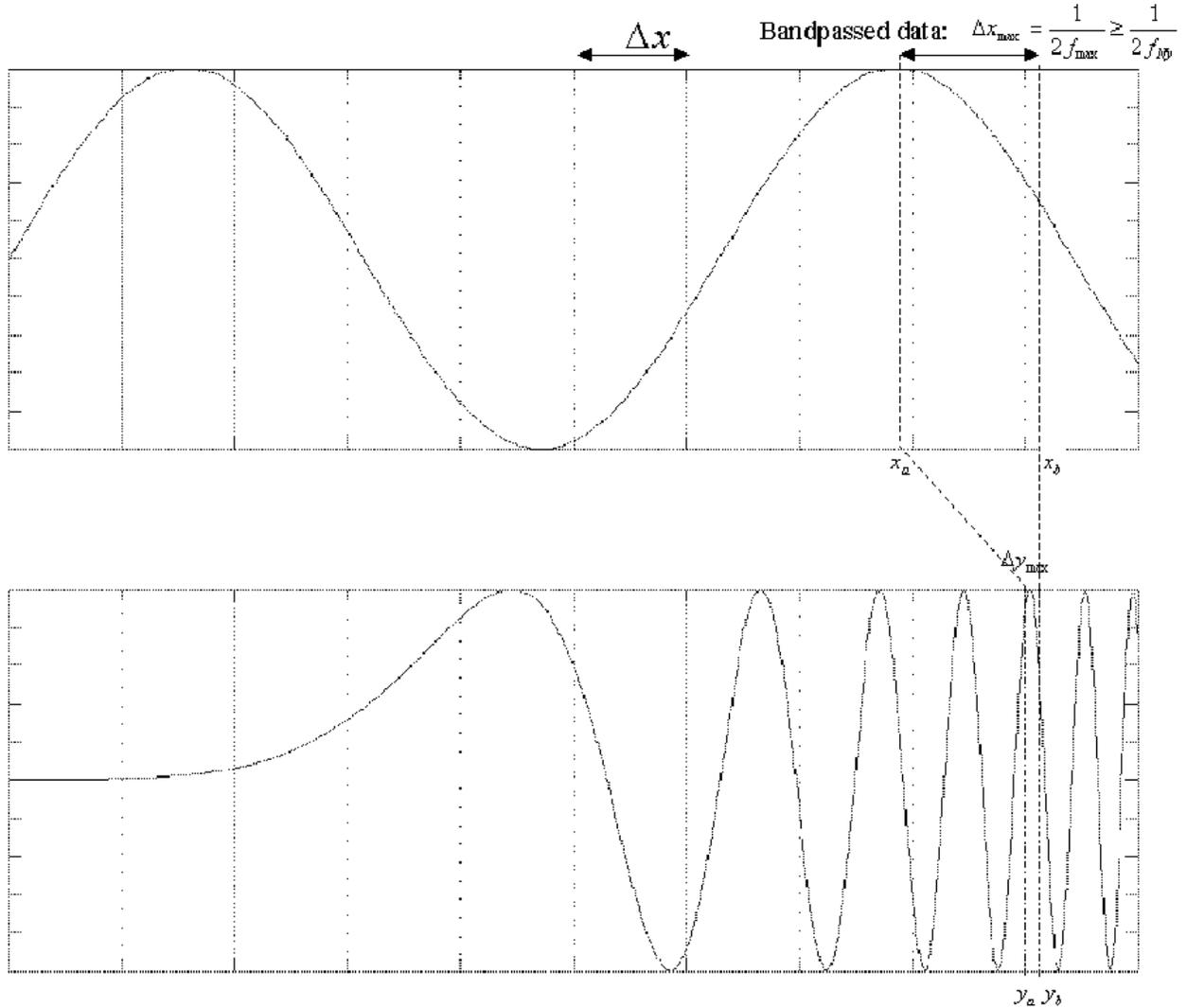


Figure 2: Illustration of how aliasing can occur while stretching: if the same sampling rate is used for the y -space (lower plot) as for the x -space (upper plot), serious aliasing will occur when transforming back to x -space. This will not happen if the sampling rate in the y -space is smaller than or equal to Δy_{\max} [nick2-strali] [NR]

2, such as

$$x_b = x_a + \Delta x_{\max} \quad (7)$$

and y_a and y_b , the images of x_a and x_b in the y space. Thus,

$$\Delta y = y_b - y_a = f(x_a + \Delta x_{\max}) - f(x_a)$$

The largest sampling rate in the y -space that will not result in aliasing is Δy_{\max} , the minimum possible value of Δy . Suppose there is a value x_m that minimizes Δy . Then,

$$\Delta y_{\max} = \left[f_{(x+\Delta x_{\max})} - f_{(x)} \right] \Big|_{x_m}$$

In particular, in the case of log-stretch, given by equation (1), if t_m plays the role of x_m from the equation above, then

$$\Delta \tau_{\max} = \left[\log \left(\frac{t + \Delta t_{\max}}{t_c} \right) - \log \left(\frac{t}{t_c} \right) \right] \Big|_{t_m} = \log \left(1 + \frac{\Delta t_{\max}}{t_m} \right) \quad (8)$$

τ_{\max} will be minimum when t_m is as large as possible, thus minimizing the expression under the logarithm. How large can t_m get? Since the length of the seismic trace is limited to a value t_{\max} ,

$$t_m = t_{\max} - \Delta t_{\max}$$

because t_m is the equivalent of x_a from eq. (7) and Fig. 2. Thus, we get

$$\Delta \tau_{\max} = \log \left(\frac{t_{\max}}{t_{\max} - \Delta t_{\max}} \right) \quad (9)$$

F-K FILTERING

As it can be seen in Fig. 3, the impulse response of the AMO computed in the log-stretch, frequency-wavenumber domain has some artifacts: high amplitude, large saddle corners. Low temporal frequencies and high spatial slopes are also present. These artifacts can be eliminated easily using a f-k filter, which is described below. Suppose we want to attenuate all spatial frequencies k that are larger than a certain threshold k_{\max} , where

$$k = \sqrt{k_x^2 + k_y^2} \quad \text{and} \quad k_{\max} = \frac{2|\omega|}{v_{\min}}, \quad (10)$$

with ω , k_x and k_y being the coordinates in the frequency-wavenumber domain (without logstretch), and v being the minimum apparent velocity of the events that we want the filtered data cube to contain. Thus, the data cube will become:

$$P_{\text{filtered}}(\omega, k_x, k_y) = \begin{cases} P(\omega, k_x, k_y) & \text{if } k \leq k_{\max} \\ e^{-\varepsilon(k-k_{\max})^2} P(\omega, k_x, k_y) & \text{if } k > k_{\max} \end{cases} \quad (11)$$

Too small an ε will result in an abrupt transition in the f-k domain, and thus ringing artifacts in the t-x domain. An ε which is too big will result in no visible filtering of the targeted artifacts. Moreover, ε depends on the choice of units and the number of samples for the m_x and m_y axes: since the exponential needs to be dimensionless, we have

$$\varepsilon = \frac{\varepsilon_0}{dk_x dk_y}$$

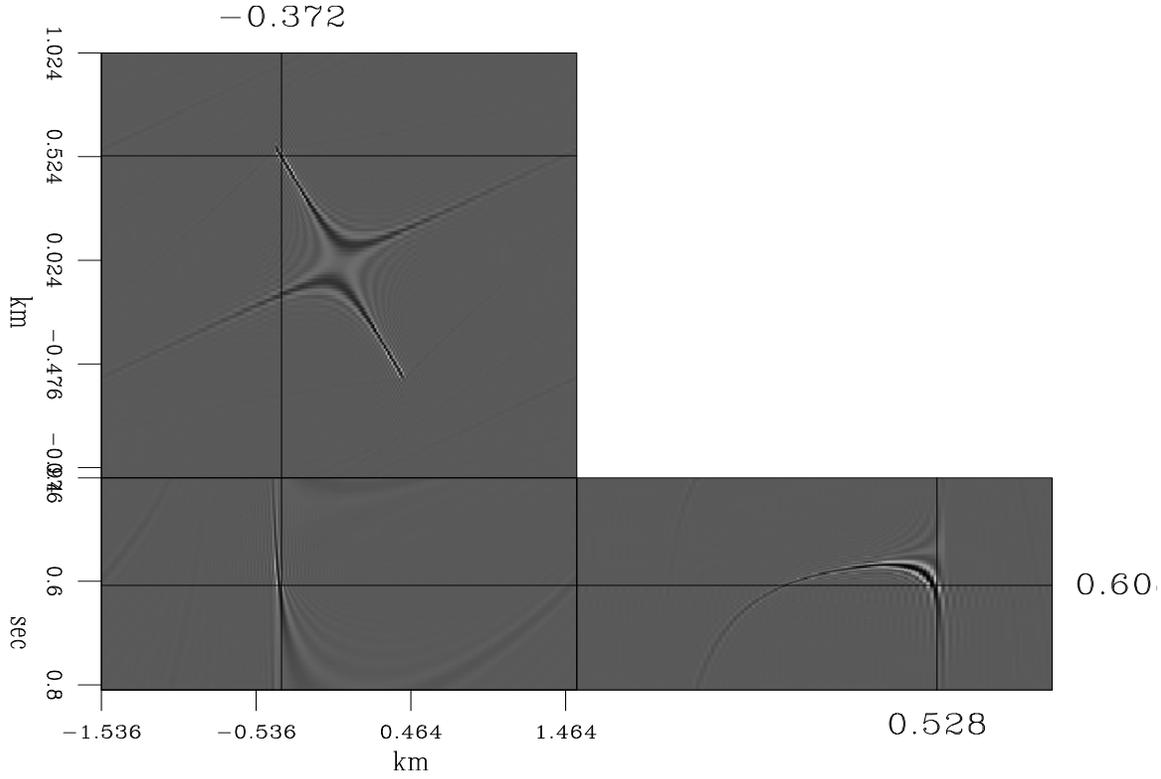


Figure 3: AMO impulse response artifacts `nick2-impresp2` [ER]

where

$$dk_x = \frac{1}{n_x d_x} \text{ and } dk_y = \frac{1}{n_y d_y}.$$

Thus, the final expression of ε is

$$\varepsilon = \varepsilon_0 n_x d_x n_y d_y, \quad (12)$$

where ε_0 is a value that is hand-picked only once, and embedded in the code. This way, we will not have to change anything at all in the code or in the parameters in order to set ε_0 , no matter what the units of the data cube may be. The result of the filtering can be seen in Fig. 4: the slices through the cube are taken at exactly the same locations as those in Fig. 3, but now the artefacts are gone.

COST-CUTTING AVENUES

The largest computational savings come from the use of FFTs for AMO, instead of slow Fourier integration necessary in the absence of log-stretch. Standard means of minimizing the CPU time and the amount of memory used to compute the AMO have also been employed. They include computing the AMO shift for only half of the elements of the cube in the complex domain, since the Fourier transform F of a real function is Hermitian:

$$F(s) = F^*(-s) \quad (13)$$

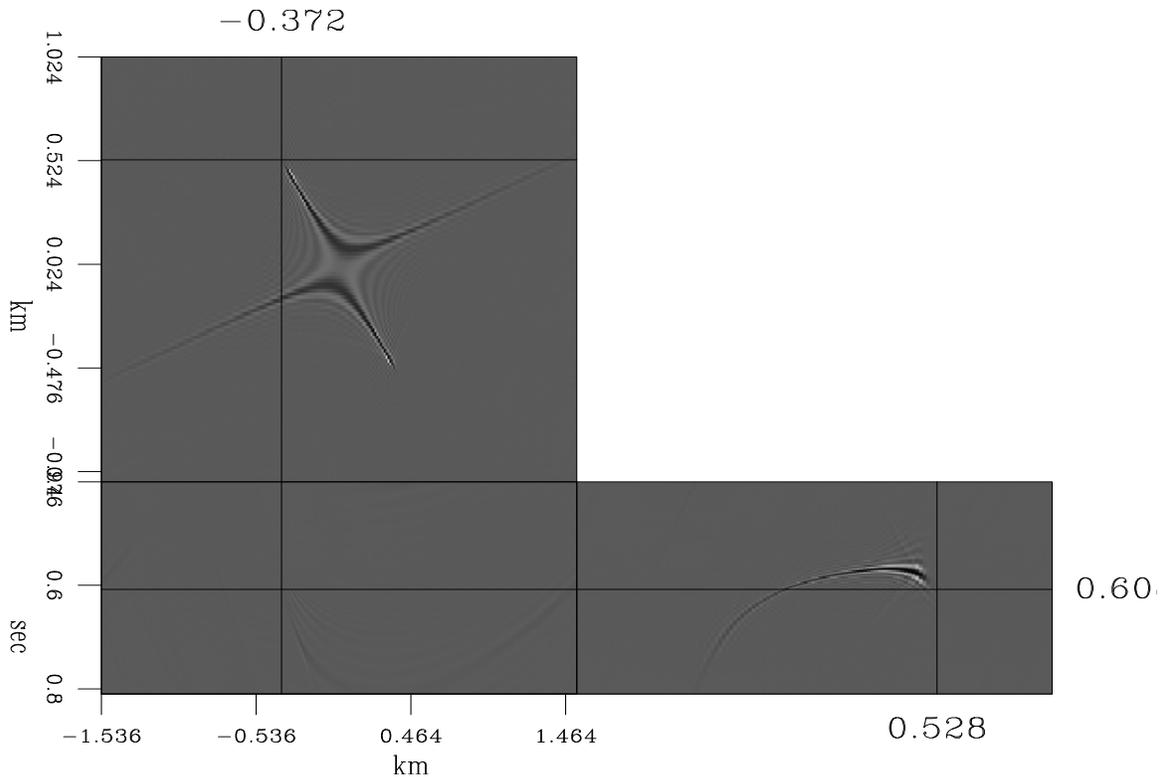


Figure 4: AMO impulse response after f-k filtering `nick2-fkfilter` [ER]

(where s denotes the frequency domain variable and the star symbol denotes the complex conjugate). Another way of reducing computational expenses was through the use of RFFTW and FFTW type Fourier Transforms (Frigo and Johnson, 1998), adaptive to hardware architecture, and taking advantage of the property stated in (13). Also, the code was divided into subroutines in such a way that some quantities were not computed unnecessarily several times when AMO was applied to more than one cube of data. Finally, shared memory parallelization with the OpenMP standard was applied to all the computationally intensive do loops in the code.

CONCLUSIONS

Azimuth moveout can be successfully implemented in the log-stretch, frequency-wavenumber domain. It is accurate, fast, and furthermore it does not have any characteristics that can result in coding difficulties.

REFERENCES

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

- Biondi, B., Fomel, S., and Chemingui, N., 1998, Azimuth moveout for 3-d prestack imaging: *Geophysics*, **63**, no. 2, 574–588.
- Black, J. L., Schleicher, K. L., and Zhang, L., 1993, True-amplitude imaging and dip moveout: *Geophysics*, **58**, 47–66.
- Bolondi, G., Loinger, E., and Rocca, F., 1982, Offset continuation of seismic sections: *Geophysical Prospecting*, **30**, 813–828.
- Frigo, M., and Johnson, S. G., 1998, Fftw: An adaptive software architecture for the fft: ICASSP Conference, Proceedings, 1381–1384.
- Gardner, G. H. F., 1991, Interpolation, crossline migration and inline depth migration of 3-d marine surveys: *SAL Annual Progress Review*, **24**, 57–69.
- Zhou, B., Mason, I. M., and Greenhalgh, S. A., 1996, An accurate formulation of log-stretch dip moveout in the frequency-wavenumber domain: *Geophysics*, **61**, no. 3, 17–23.

Short Note

On asymmetric Stolt residual migration for imaging under salt edges

Daniel Rosales and Biondo Biondi ¹

INTRODUCTION

Imaging under salt remains a problem for the oil industry (Muerdter et al., 1996; Prucha et al., 1998) due to illumination problems and poor velocity resolution under the salt body. Residual Stolt migration is a proved technique for updating the image and the velocity field (Sava, 1999), because of that we believe that an asymmetric residual migration operator can help in updating the image just below the salt body.

Rosales et al. (2001) introduced Stolt residual migration for converted waves. The new operator has the property of correcting the source-reflector and reflector-receiver leg independently. This property can be extended for imaging complex single mode (PP) data (e.g. under salt imaging).

We start with a review of the theory already discussed by Rosales et al. (2001) in order to handle separately the source and receiver legs in single mode data. We present a potential frame for imaging update under salt edges, as well as future plans on this area.

THEORY REVIEW

Rosales et al. (2001) presented residual pre-stack Stolt migration for converted waves, this new operator is asymmetric and this asymmetry property can be used for imaging under salt edges, by keeping unchanged one travel-time leg and modifying the other.

The asymmetric pre-stack Stolt residual operator can be written as:

$$k_{zm} = \frac{1}{2}\sqrt{\rho_s \kappa_0^2 - k_s^2} + \frac{1}{2}\sqrt{\rho_g \gamma_0^2 \kappa_0^2 - k_g^2} \quad (1)$$

where

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

¹email: daniel@sep.stanford.edu, biondo@sep.stanford.edu

$\rho_s = \frac{v_{s0}}{v_{sm}}$, $\rho_g = \frac{v_{g0}}{v_{gm}}$, and $\gamma_0 = \frac{v_{s0}}{v_{g0}}$; the subscripts s and g denote for source and geophone, respectively.

The definition of the sources and receivers position is important since it explains the meaning of the three parameters used for the residual migration. If we define the source position along the fastest path, the value of γ_0 will be bigger than one. The contrary would happen if we define the geophone position along the fastest path.

METHODOLOGY

Figure 1 illustrates the main idea of this work. We are using a synthetic velocity model with a simple reflector geometry but a complicated salt structure that challenges the image under the salt body (Prucha et al., 2001). We define two groups of rays for the same reflector point, group A travels through the sediments and group B travels through the salt body. Since imaging under salt body is relative complicated (Prucha et al., 2001), we are going to modify and update our image of the reflections points that have, at least, one group of rays traveling through the salt body.

The imaging update is possible by just modifying the rays that goes through the salt body, process that is possible by fixing one of the parameters (ρ_s) in equation (1) to 1, and by varying the other two (ρ_p and γ_0).

It is possible to observe in Figure 2, the different effect that this procedure will have for different reflectors points. Figure 2a presents a reflector point away from the salt, for this reflector we should have no correction at all, since there are no rays going through the salt. Figure 2b presents a reflector point with few rays going through the salt, the correction should be small for this reflectors. Figure 2c presents a reflector point with a significant number of rays going through the salt body, these kind of reflectors are generally difficult to image and for those we expect to improve the quality of the image with the asymmetric pre-stack Stolt residual migration operator.

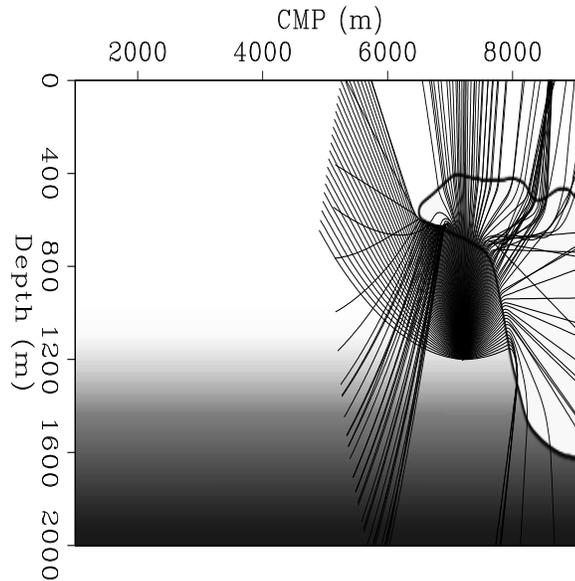
It is also possible to separate the image in sections where we need to apply correction (Figure 2c), and sections where we do not need to apply the correction (Figure 2a). Therefore, we need to be able to interpolate our final image from those two separate sections.

SUMMARY AND FUTURE WORK

This idea is a work in progress. In this paper we presented how we can use a tool developed for updating converted waves images for imaging complicated geological regions with single mode data.

We are planning to apply this methodology to the velocity model presented in this paper. We would also like to apply to a real data set, that presents interest challenging problems under salt edges.

Figure 1: Main problem: Two different groups of rays, group *A* doesn't go through the salt body; group *B* goes through the salt body.
daniel2-hwt2c [ER]



REFERENCES

- Muerdter, D. R., Lindsay, R. O., and Ratcliff, D. W., 1996, Imaging under the edges of salt sheets: a raytracing study: 66th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 578–580.
- Prucha, M. L., Clapp, R. G., and Biondi, B. L., 1998, Imaging under the edges of salt bodies: Analysis of an Elf North Sea dataset: SEP-97, 35–44.
- Prucha, M. L., Clapp, R. G., and Biondi, B. L., 2001, Imaging under salt edges: A regularized least-squares inversion scheme: SEP-108, 91–104.
- 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.

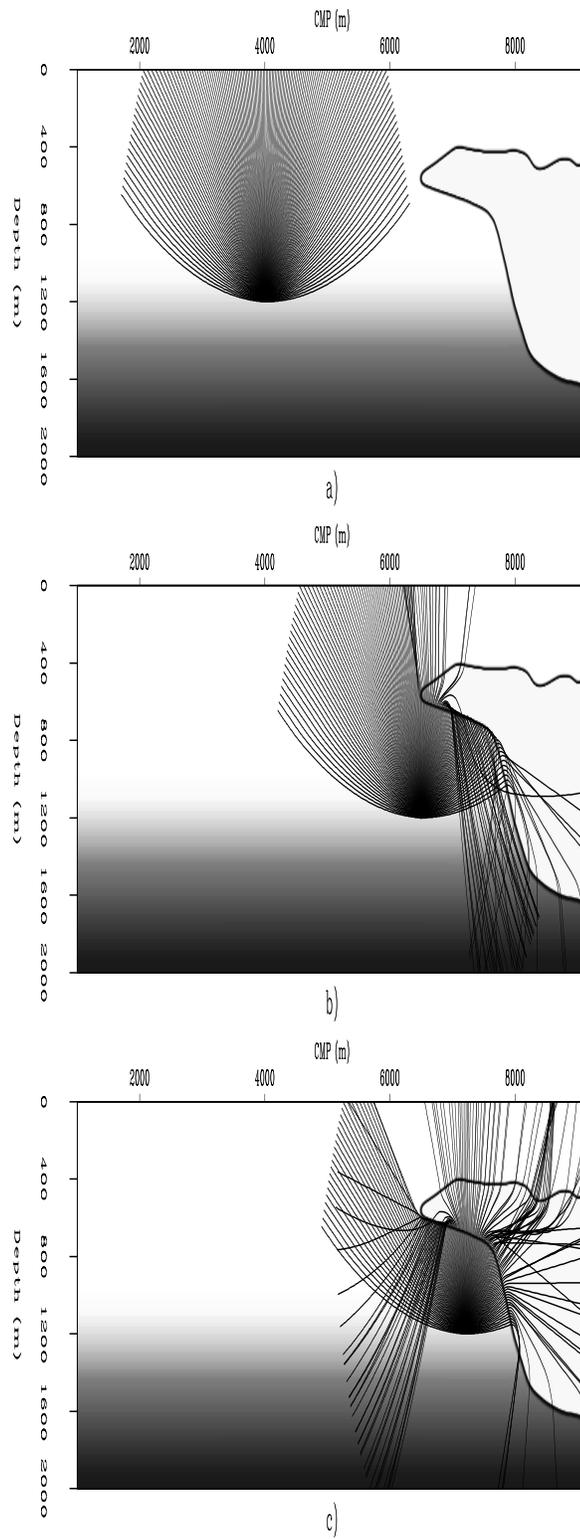


Figure 2: Different ray tracing for different reflector position. a) A reflector position not affected by the salt body; b) A reflector position partially affected by the salt body; c) A reflector position totally affected by the salt body. [daniel2-hwt2d2](#) [ER,M]

Data dependent parameterization and covariance calculation for inversion of focusing operators

*Barbara E. Cox*¹

ABSTRACT

The Common Focus Point (CFP) method makes it possible to convert conventional two way traveltimes to focusing operators that represent one way traveltimes. These focusing operators are inverted to obtain a velocity model. The under-determined nature of the inversion problem is addressed by the data dependent adjustment of the parameterization by means of the a posteriori covariance. This results in an efficient, non-laborious algorithm producing well determined inversion problems. The required a posteriori covariance is normally explicitly solved in the explicit matrix calculation during optimization. However, these explicit calculations are not feasible in larger problems. Fortunately, algorithms have been proposed to extract the a posteriori covariance from the more efficient approximate matrix inversion algorithms that are available.

INTRODUCTION

The quality of a seismic image is highly dependent on the accuracy of the velocity model of the subsurface. The CFP method (Berkhout, 1997a,b; Thorbecke, 1997) has proven to be an appropriate tool to estimate this velocity model, because the inversion of one way traveltime data (focusing operators) generated by this method is inherently simpler than the inversion of conventional two way traveltime data (Kabir and Verschuur, 2000; Hegge, 2000).

In this report, the inversion of focusing operators is also used to obtain a velocity model. As in all geophysical inversion problems, the under-determined nature is a problem that should be faced. Normally, this is handled by two methods; 1) choosing a well determined parameterization (global parameterization by user defined layers), or 2) regularizing the optimization (including a priori information and model covariance). A drawback of the first method is that it puts a constraint on the result and it can be laborious. Moreover, incorrect initial parameterization can lead to slow or non-convergence. A drawback of the second method is that the problem of over-parameterization is still not solved (e.g. when a regular grid is used). This might cause problems when the inversion problem is regularized; the model parameters need different levels of regularization, as certain regions in the model might consistently be more under-determined. In the data dependent parameterization shown in this report both methods will be combined; adjustment of the parameterization is based on the covariance after opti-

¹email: B.E.Cox@ctg.tudelft.nl

mization. In the explicit least squares or singular value decomposition (SVD) optimization methods, the a posteriori covariance is generated during optimization. However, in more practical approximate optimization methods, obtaining the covariance becomes more difficult.

In this report, the concept of the CFP method and the focusing operator updating will be explained in the first section. Next, the inversion of the focusing operators, and the data dependent parameterization will be considered. The method will be applied to a synthetic example. The last part will compare the different approximate inversion algorithms for obtaining the inverse and the a posteriori covariance. Finally, some conclusions and plans for future work will be presented.

CFP METHOD

Focusing operators

For each point in the subsurface, seismic migration can be written in terms of two focusing steps: focusing in detection followed by focusing in emission. The focusing is done by a focusing operator, which can be considered as a one-way seismic response from one point in the subsurface to locations at the surface (Fig.1).

The first focusing step is performed by time-convolving a shot record with the focusing operator of the point under consideration. Next, all traces in the obtained move-out corrected shot record are summed (consider *shot 1* in Fig.1). When this first focusing step is applied to all shots, each shot record is transformed into a single trace by the focusing operator (consider *shot 1,2-n* in Fig.1). Together those traces define the so-called common focus point gather (CFP-gather). Each trace in the CFP gather is positioned at the source position of its corresponding shot record. One event in the CFP-gather is the focus point response. If the velocity model is correct, the focus point response and the time-reversed focusing operator (= Green's function) have equal traveltimes: principle of equal traveltimes (Fig.1).

Note that the first focusing step "transforms" conventional two way data to one way CFP data. The second focusing step is performed by applying the focusing operator again. This procedure transforms the focus point response into the seismic image of the subsurface grid point under consideration (CFP-stacking). In this report only the first step is considered, because this is the required step for obtaining focusing operators as will be shown in the next section.

Concept of focusing operator updating

The principle of equal traveltimes formulates that the time-reversed focusing operator (based on the underlying velocity model) and the focus point response (based on the underlying seismic measurements) must have equal traveltimes. In the situation of velocity errors both the operator and the response contain an error. The error in the operator has an opposite effect on the error in the response. Therefore, the exact operator is situated somewhere between the wrong

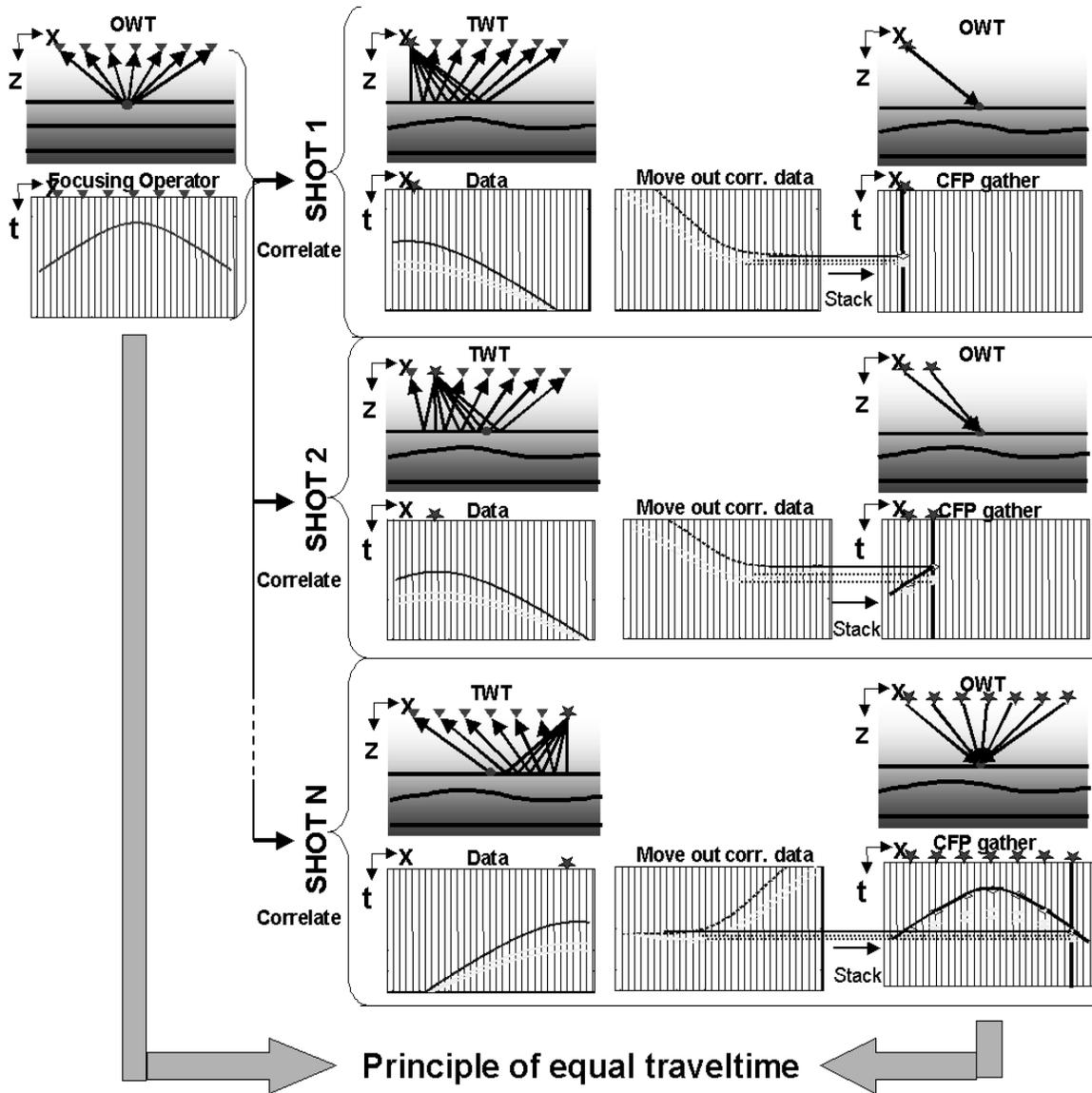


Figure 1: CFP method explained schematically, by both ray pictures and time sections; a modeled focusing operator (one way traveltime (owt)) is applied to a shot (two way traveltime data(twt)), resulting in move out corrected data. The stack of these data result in one trace of a CFP gather (owt). This is done for all (1,2.....N) shots, generating a complete CFP gather. When a correct focusing operator is applied, the time-reversed focusing operators and the focus point response (one event in CFP gathers) have equal travelttime. barbara-fop [NR]

operator and the wrong response, depending on the velocity error and the local dip. A data driven update procedure (Berkhout, 1997a; Thorbecke, 1997; Bolte and Verschuur, 1998) for the focusing operator is as follows (Fig.2):

- Cross-correlate each trace of the time-reversed focusing operator with the corresponding trace of the CFP-gather. The result is a correlation panel, which contains stretch-free move-out corrected data around zero time;
- Update each trace of the focusing operator with half the corresponding time-shift in the correlation panel.

Finally, the correlation panel will show a flat event at zero time, indicating that the principle of equal traveltimes is satisfied (Fig.2). This procedure will result in exact focusing operators, indicating the response from one point in the subsurface to acquisition locations at the surface. Note that at this stage, no velocity model is involved yet. By the inversion of the traveltimes of these focusing operators a velocity model of the subsurface may be derived .

FOCUSING OPERATOR INVERSION

The CFP velocity estimation that is described in this report is based on a separated inversion method, in which first the focusing operators are obtained from the conventional data (focusing operator updating as described in the previous section), and after that the velocity model is estimated by inverting these focusing operators. In the model, both the velocity and the exact locations of the focus points are unknown. The inversion of the focusing operators should result in a velocity model (that represents the velocity structure of the subsurface) and the exact location of the focus points (that corresponds to reflection energy). This inversion process is carried out with traveltimes tomography.

Traveltimes tomography

Traveltimes tomography generates a velocity model of the subsurface from the observed traveltimes response through the subsurface. In focusing operator inversion, the traveltimes response is the traveltimes of the focusing operator. Traveltimes tomography contains three main steps: parameterization of the velocity in the subsurface; forward modeling through the subsurface model to obtain modeled focusing operators; and optimization in which the difference between the traveltimes of the observed and the modeled focusing operators is used to optimize the velocity model and the focus points positions. If the traveltimes of the observed and the modeled focusing operator correspond, an accurate velocity model is obtained. The strategy for focusing operator inversion is presented in Figure 3.

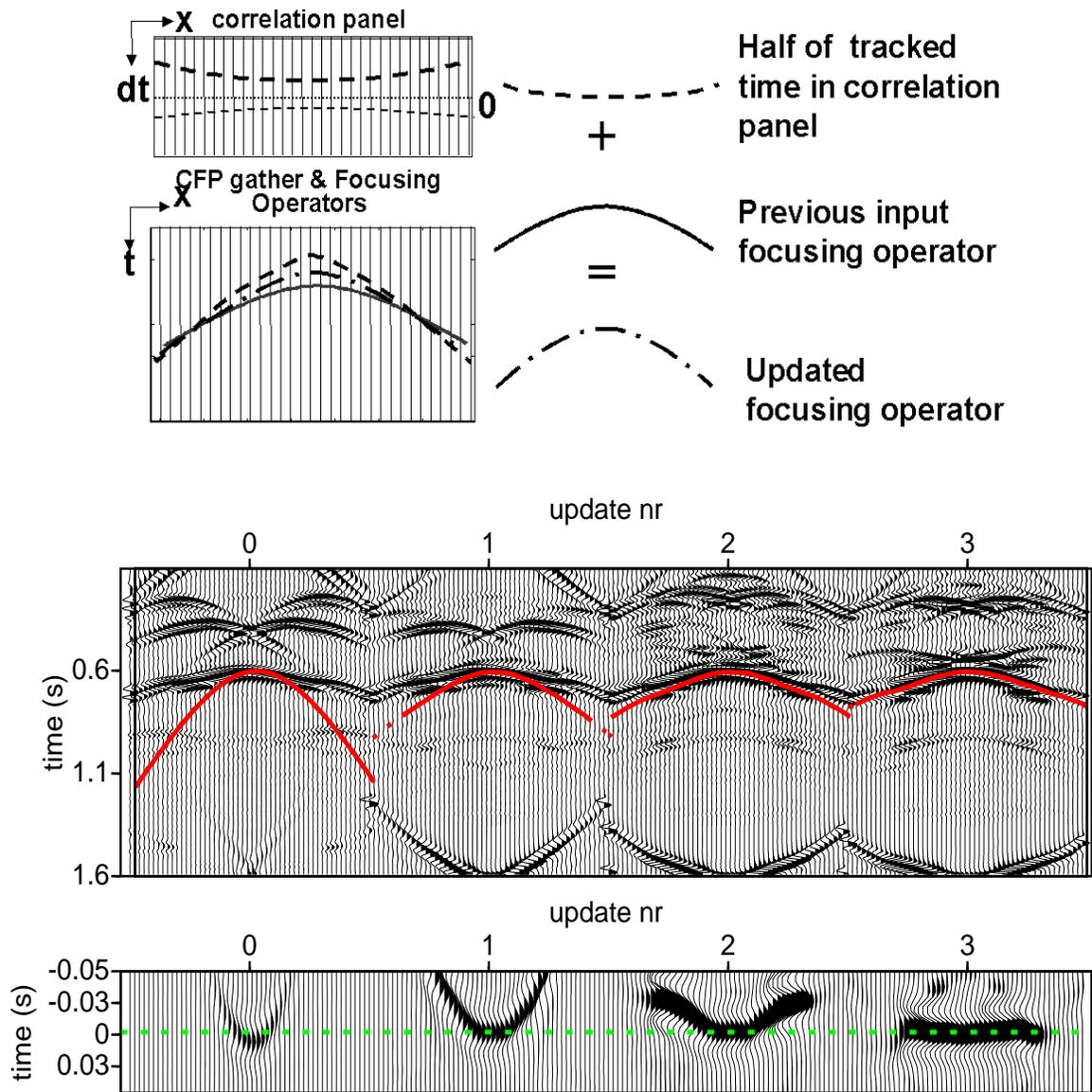
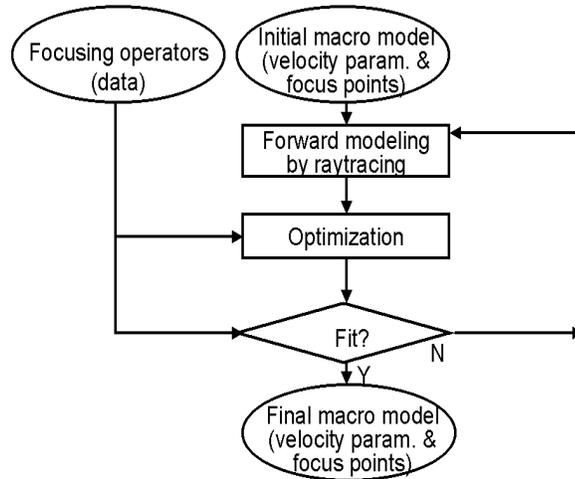


Figure 2: Focusing operator updating. *Top:* Adding half the time shift (dashed line) in correlation panel to the initial focusing operator (solid line in upper plot) results in an updated focusing operator (dash-dotted line in lower plot). Dashed line in lower plot represents the focus point response in the CFP gather. *Bottom:* 4 iterations of focusing operator updating; upper plots show CFP gather and focusing operator (gray line); lower plots show correlation panels. Note that after 4 iterations this panel is flat due to the principle of equal traveltimes (after Bolte (1998)). [barbara-fopup] [NR]

Figure 3: Strategy flowchart of inversion of focusing operators
 [barbara-flow] [NR]



Parameterization

There are several ways to describe the velocity model by a set of parameters. A layered parameterization described by global basis functions can be used, i.e. the velocity is estimated for regions and layers in the subsurface defined by the user. However, this parameterization constrains the possible range of solutions. Parameterization based on local basis functions is a more flexible description of the velocity model. The model can be described by cells or grid-points at which the values of the parameters (i.e. the values of the velocity field) are defined.

In the algorithm for tomographic inversion of focusing operators, Delaunay triangulations are used to construct cells (triangles) between grid-points (Fig.4a,b). The velocity is defined at the grid-points and the velocity within the triangles is calculated by a linear interpolation between the three grid-points defining each triangle (Fig.4b). In this way, every point can adopt an optimum velocity, and any kind of subsurface can be described. The focus points are defined independently of the velocity points. The focus points are related to positions where reflection energy is available. By parameterizing them independently of the velocity grid-points, the velocity changes are not dependent on, or constrained to, the reflectors.

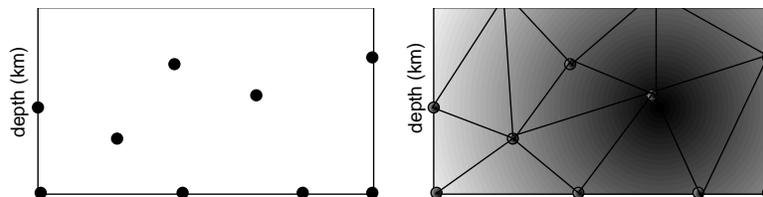


Figure 4: Parameterization of the velocity model. (a) The velocity is defined at grid-points (black dots) in the subsurface. (b) The grid-points are connected by Delaunay triangulation. The velocity within the triangles is calculated by linear interpolation.
 [barbara-param] [NR]

Forward Modeling

Forward modeling is done by ray tracing in the initial velocity model. An imaginary shot (focus point) is positioned in the subsurface and rays are traced towards receivers at the surface. In this way modeled focusing operators are obtained. The differences between the traveltimes of these operators and the real focusing operators are used to optimize the velocities in the model and the locations of the focus points. Moreover, the traced rays are needed to define the relation between the traveltime data and the model parameters (matrix \mathbf{A} , in the next subsection).

Optimization

Optimization is done by calculating the updates of the model parameters by means of the difference between the traveltimes of the modeled and the observed focusing operators. The relation between the traveltime data and the model parameters is assumed to be linear:

$$\Delta \mathbf{d} = \mathbf{A} \Delta \mathbf{m}, \quad (1)$$

where $\Delta \mathbf{d}$ contains the difference in traveltimes, $\Delta \mathbf{m}$ contains the update in model parameters, and \mathbf{A} describes the linear relation between these traveltimes and the model parameters. It might be very difficult to solve for the model vector $\Delta \mathbf{m}$ in terms of the data vector $\Delta \mathbf{d}$, as matrix \mathbf{A} generally is not invertible. However, the generalized matrix inverses can be obtained by several methods. The generalized inverse obtained by these methods is defined by a dagger:

$$\Delta \mathbf{m} = \mathbf{A}^\dagger \Delta \mathbf{d}, \quad (2)$$

When a least squares method is chosen to obtain a generalized inverse, an update of $\Delta \mathbf{m}$ is calculated by:

$$\Delta \mathbf{m} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \Delta \mathbf{d}, \quad (3)$$

where $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ is the generalized inverse. Another way to obtain the generalized inverse is for example by singular value decomposition (SVD). Berryman (2001a) gives a profound overview of the available methods and their capacities. Several methods of approximating the generalized inverse will be presented in one of the next sections.

DATA DEPENDENT PARAMETERIZATION

As inversion problems are generally under-determined, they should be stabilized. Stabilization methods can be divided into two groups:

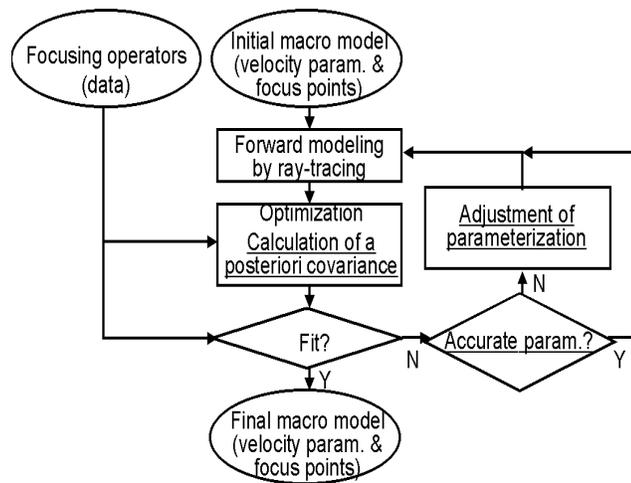
- Regularization of optimization;
- Adjustment of parameterization;

The first method can be performed by both regularization criteria (e.g. damping of the least squares method, or truncated SVD) and model regularization (e.g. including a priori information or conditional covariance matrices in the optimization). A drawback of this method, however, is that the problem of over-parameterization is still not solved, which might cause problems when the inversion problem is regularized. The second method, the adjustment of parameterization, can for example be done by using a global parameterization. However, as mentioned before, this parameterization constrains the possible range of solutions, and incorrect initial parameterization can lead to slow or non-convergence.

In the inversion of focusing operators an attempt is made to combine both stabilizing methods; the adjustment of parameterization is based on the covariance after optimization; the a posteriori covariance. Model parameters that have a high variance are removed, and in regions containing low variance parameters, extra parameters can be added. In this way, the inversion problem is well determined and over-parameterization is avoided. Moreover, all the available information within the data will be translated to the model, as the parameterization adapts to the data.

If the data is used to determine the parameterization, as a consequence the obtained parameterization tells something about the data. A coarse parameterization exposes for example that more focus points should be positioned in that region (of course this is only possible if data related to the focus point is available). On the other hand, when updated focus points are positioned very close to each other, it might be wise to remove/shift them, as they will only duplicate information.

Figure 5: Strategy flowchart for inversion if focusing operators with data dependent parameterization. The extra steps w.r.t. Figure 3 are underlined. barbara-flowap [NR]



A posteriori covariance

The a posteriori covariance is an important tool to evaluate the solution of the optimization. Therefore it is useful to use it as a criteria for the adjustment of the parameterization. The covariance matrix can be calculated during optimization by (Menke, 1984):

$$\mathbf{C} = \mathbf{A}^\dagger (\mathbf{A}^\dagger)^T, \quad (4)$$

When least squares optimization (shown in equation (3)) is used, the first term on the right hand side corresponds to the covariance matrix, so

$$\mathbf{C} = (\mathbf{A}^T \mathbf{A})^{-1}. \quad (5)$$

The covariance matrix contains M diagonal elements that correspond to the (M) variances of the model parameters. These elements are used to adjust the parameterization after optimization (Fig.5).

In Figure 6 an example of the use of a posteriori covariance for data dependent parameterization is shown. For the velocity model shown in Figure 6a the traveltimes data corresponding to the rays in Figure 6b are available. When the a posteriori covariance matrix is calculated and the diagonal elements of the matrix (variances) are plotted for each grid-point, this results in an image as shown in Figure 6c. In this image the dark values indicate low variance and light values indicate high variance. As mentioned before, high variance grid-points are removed, and in regions containing low variance grid-points extra parameters can be added. This finally results in the data dependent parameterization shown in Figure 6d.

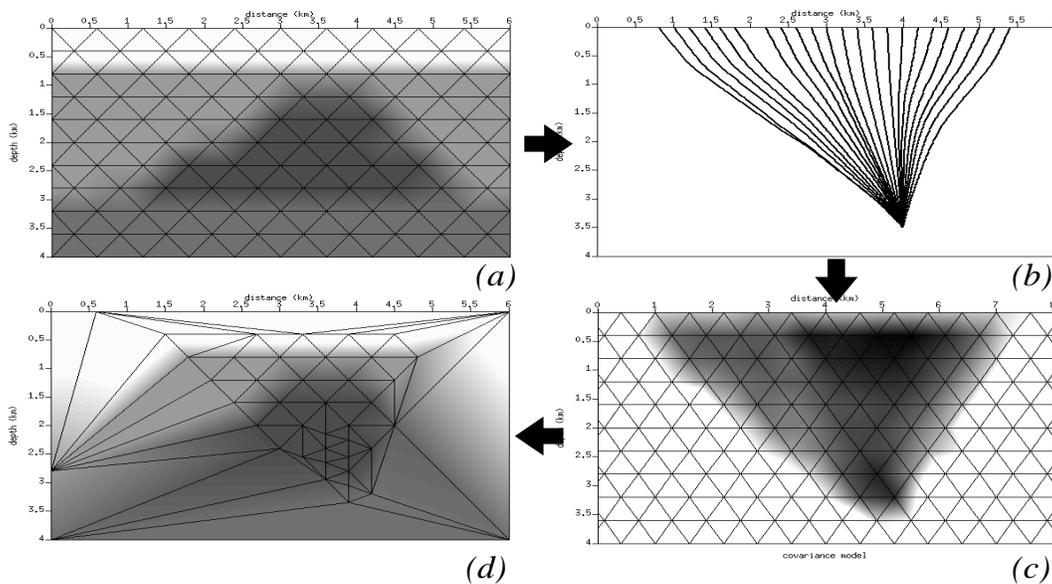


Figure 6: Example of data dependent parameterization. (a) Delaunay triangulation in model defined by a regular grid. (b) Ray-paths from a focus point at position (4.0km; 3.5km) to the surface. (c) Variances at the grid-points in the optimization with ray-paths in (b). Dark values indicate low variance, light values indicate high variance. (d) Adjusted parameterization based on the variances. barbara-datadep [NR]

A synthetic example

The algorithm for tomographic inversion of focusing operators with data dependent parameterization is applied to a synthetic 2D case. The model contains a salt dome, fault structures

beneath the dome, a complex turbidite velocity structure with low velocities, and lateral and vertical velocity gradients within the layers (Fig.7a). This is the "real" velocity model of the subsurface, in which (in this synthetic case) the focusing operators are generated (Fig.7b). The positions of the focal points are represented as black dots in Fig.7a. An initial velocity model and an initial estimate of the focal point positions are needed to start the updating procedure (Fig.7c). From these initial locations focusing operators are modeled in the initial velocity model (Fig.7d). Note that for display purposes, a wavelet has been assigned to each operator, and the resulting wave fields have been shown. In the inversion algorithm, only the traveltimes are used. By the traveltime difference between the modeled operators and the real operators, the velocity model and focal point positions are updated. In each iteration, the parameterization will be adjusted and the velocity model and focal point positions will be updated.

The final result after 6 iterations is shown in Figure 8 (displayed in terms of squared-slowness). Note that at locations with high data information (i.e. originally low variances), the density of the grid-points is high. The corresponding velocity model of the final result is also shown in Figure 8. The positions of the focal points are similar to the real focal point positions in Figure 7a. The velocity model also resembles the real model, although it is a smooth solution of the real case. However, the upper two layers (water layer, and layer below) are not well resolved, which is indicated by the higher velocity, and the downward shifted focus points. This artifact is caused by the severe ray distortion originated in the salt dome. When this velocity model is used for migration (Fig.8), most reflectors are clearly visible. Nevertheless, the inaccuracy of the upper two layers (water layer, and layer below) is also visible in this migration. In particular above the salt dome, where the reflectors are shifted downward.

APPROXIMATE COVARIANCE CALCULATION

For small scale problems, generalized inverses can be calculated easily by explicit matrix calculation during least squares optimization, or during SVD. However, for larger problems this becomes impracticable. Several methods for approximation of the generalized inverse have been developed to circumvent this problem. In these methods no explicit matrix calculation is required. However, as the generalized inverse can not be solved explicitly, neither can the covariance matrix be provided by these methods. Several methods for obtaining the covariance matrix from the generalized inverse approximation are available. The approximation methods for the inverse and the covariance can be divided in two groups: iterative methods and model space weighting methods. Both approaches are considered in this section.

Iterative methods

The iterative methods for obtaining the generalized inverse are based on the SVD method. The available methods are among others Conjugate Gradient (CG) (Hestenes and Stiefel, 1952) and LSQR (Paige and Saunders, 1982). Berryman (2001b) provides a complete overview of the iterative methods, and an analysis of their capacities. All the iterative methods lead directly to approximations of the generalized inverse \mathbf{A}^\dagger . In general, carrying these iterative processes

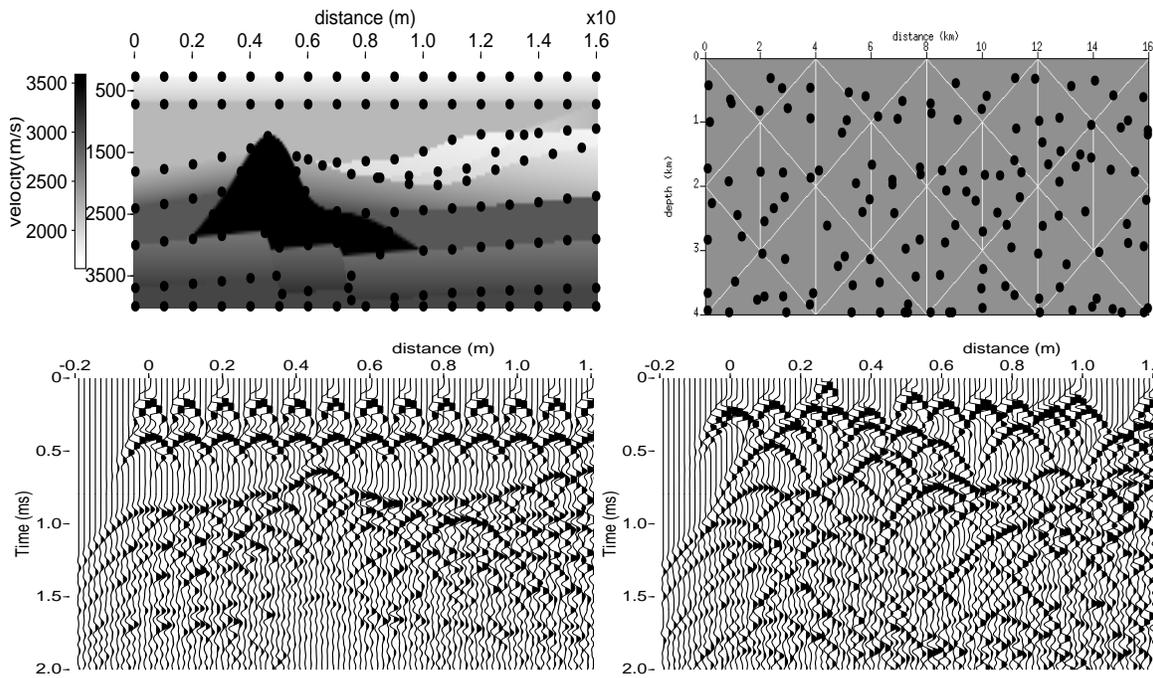


Figure 7: Velocity models of a synthetic case. Black dots represent focal points. (a) Real velocity model and focal points (b) Real focusing operators (modeled in (a), using a wavelet for display); (c) Initial model (squared slowness) and initial focal points; (d) Focusing operators modeled in (c) `barbara-synth` [NR]

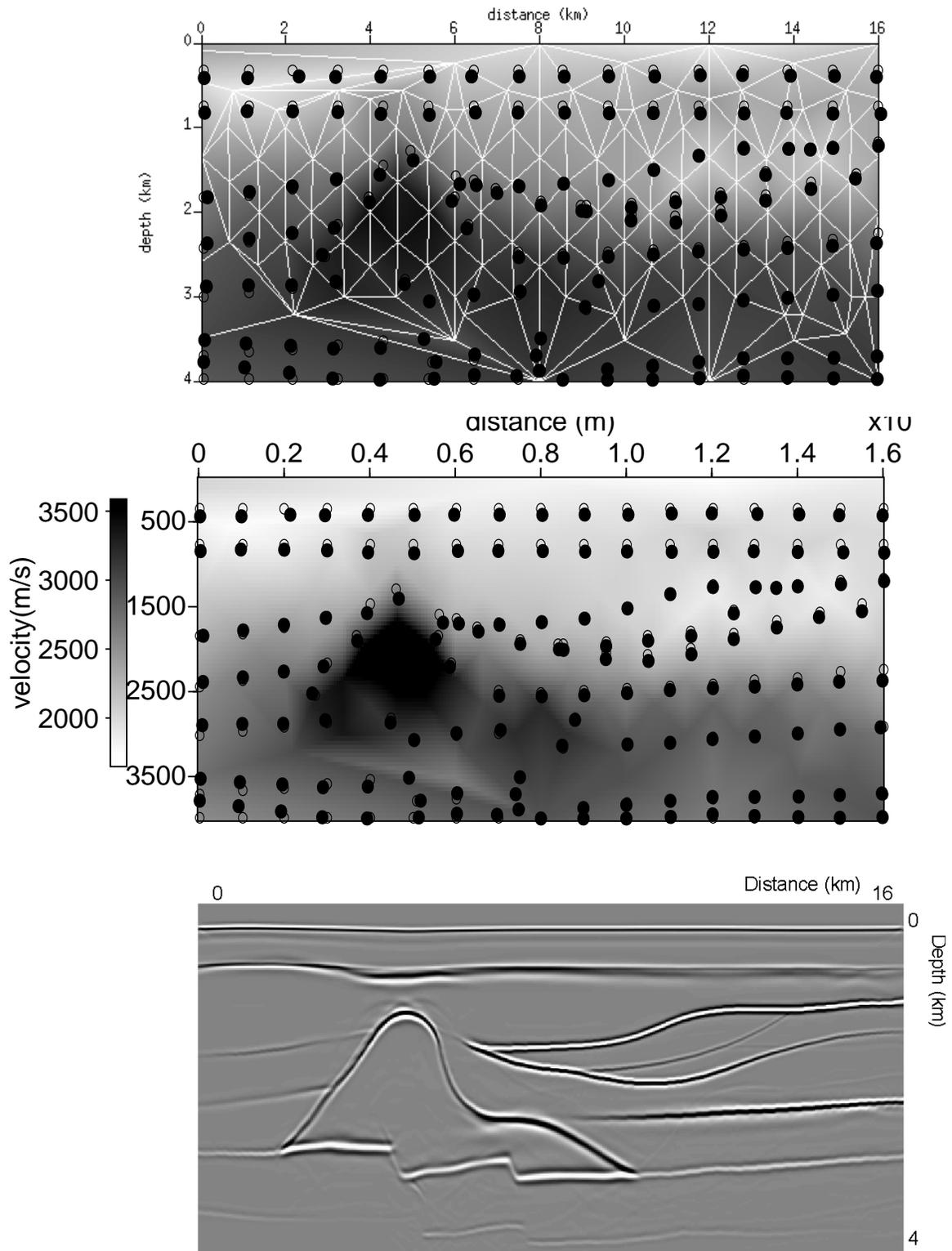


Figure 8: *Top*: Final (sixth) update of model (squared slowness) and focal points; open dots: real focus point locations, black dots: updated focus point location; *Middle*: Final update of model (velocity) and focal points. *Bottom*: Migration with velocity model displayed in figure above. `barbara-res` [NR]

to completion will produce a result closely approximating \mathbf{A}^\dagger .

Yao et al. (1999) and Berryman (2001b) provide methods for calculating the a posteriori covariance for iterative methods based on SVD. For a thorough description of the methods I refer to these papers. The principle is as follows; each iteration provides one extra vector in the solution space. After K iterations these K vectors are used for calculating the a posteriori covariance. As a consequence, if more iterations are performed, not only the approximated inverse but also the a posteriori covariance becomes more accurate. Figure 9b,c shows a posteriori covariance matrices for $M = 33$ model parameters obtained by respectively CG and LSQR. Both matrices resemble the real a posteriori covariance matrix (Fig.9a). Note that both methods compute the complete covariance matrix, so also the non diagonal elements.

Model space weighting methods

Another approach of approximating the inverse and the a posteriori covariance matrix is provided by the work of Claerbout and Nichols (1994). Rather than trying to solve the explicit inverse problem the adjoint of \mathbf{A} is applied to the data, and a diagonal operator \mathbf{W}_m^2 is constructed, such that

$$\Delta \mathbf{m} = \mathbf{W}_m^2 \mathbf{A}^T \Delta \mathbf{d}, \quad (6)$$

in which \mathbf{W}_m^2 is related to the model space weighting. This term is needed to obtain the correct physical units. Note that this term also corresponds to the a posteriori covariance. The most significant properties of the model space weighting term can be explored by measuring its effects when applied to a reference model $\Delta \mathbf{m}_{\text{ref}}$. For example, if \mathbf{W}_m^2 is given by

$$\mathbf{W}_m^2 = \frac{\text{diag}(\Delta \mathbf{m}_{\text{ref}})}{\text{diag}(\mathbf{A}^T \mathbf{A} \Delta \mathbf{m}_{\text{ref}})} \approx (\mathbf{A}^T \mathbf{A})^{-1}, \quad (7)$$

will have the same units as \mathbf{A}^\dagger , and the properties of this term can be explored. Two choices of $\Delta \mathbf{m}_{\text{ref}}$ will be shown here (Biondi, 1998):

- $\Delta \mathbf{m}_{\text{ref}}$ = unity vector
- $\Delta \mathbf{m}_{\text{ref}}$ = unit vector e^k in k direction in m -D ($k = 1, 2, \dots, m$)

The first choice results in the sum of the elements in each column of \mathbf{A} in the denominator. The second choice results in the sum of the square of the elements in each column of \mathbf{A} in the denominator. The obtained vector forms the diagonal in the approximated $(\mathbf{A}^T \mathbf{A})^{-1}$. Figure 9d,e shows a posteriori covariance matrices for $M = 33$ model parameters obtained by respectively the "sum" approximation and the "diagonal" approximation. Note that only the diagonal elements are calculated. These elements resemble the diagonal of the real a posteriori covariance matrix (Fig.9a).

Whether the iterative methods, or the model space weighting methods should be used for calculation of the covariance matrix depends also on the algorithm that is used for calculating the approximate inverse. In general, iterative methods are used for tomography, and the vectors

needed for calculation of the covariance are obtained without any extra costs. The calculated covariance matrix will reflect the status of the inversion after the performed number of iterations. However, a drawback of these methods is that the solution space vectors need to be saved after each iteration. Moreover, not performing enough iterations might result in strange covariance matrix. Nevertheless, this matrix will still reflect the status of your inversion after the performed number of iterations. The model weight methods are very cheap, and give accurate results of the diagonal elements of the covariance. Nevertheless, the non-diagonal elements are not calculated. However, as for the re-parameterization only this diagonal is considered, these methods might be sufficient. When the covariance between the different parameters is going to play a role in the inversion too, the full covariance matrix obtained from the iterative methods should be used.

CONCLUSIONS

The proposed algorithm for inversion of focusing operators results in an accurate update of both velocities and focus point locations. The data dependent parameterization is an efficient, non-laborious algorithm resulting in well determined inversion problems. The a posteriori covariance that is needed in the algorithm can also be extracted from the approximate matrix inversion algorithms that are available.

FUTURE WORK

Although in the proposed method all the available information in the data can be translated to a model, this might still not result in a clear image of the subsurface. This can for example be caused by poor data quality or missing data in crucial regions. Therefore, it should always be possible to include a priori information about the subsurface into the inversion. Including a priori velocity information is already possible, however, it might also be necessary to include other information criteria in the inversion scheme.

ACKNOWLEDGMENTS

I would like to thank the SMAART consortium for making my visit to SEP possible, and the people of the CWP project for their Delaunay and ray-tracing software, which formed a base for the developed algorithm described in this report.

REFERENCES

Berkhout, A. J., 1997a, Pushing the limits of seismic imaging, part I: Prestack migration in terms of double dynamic focusing: *Geophysics*, **62**, no. 3, 937–953.

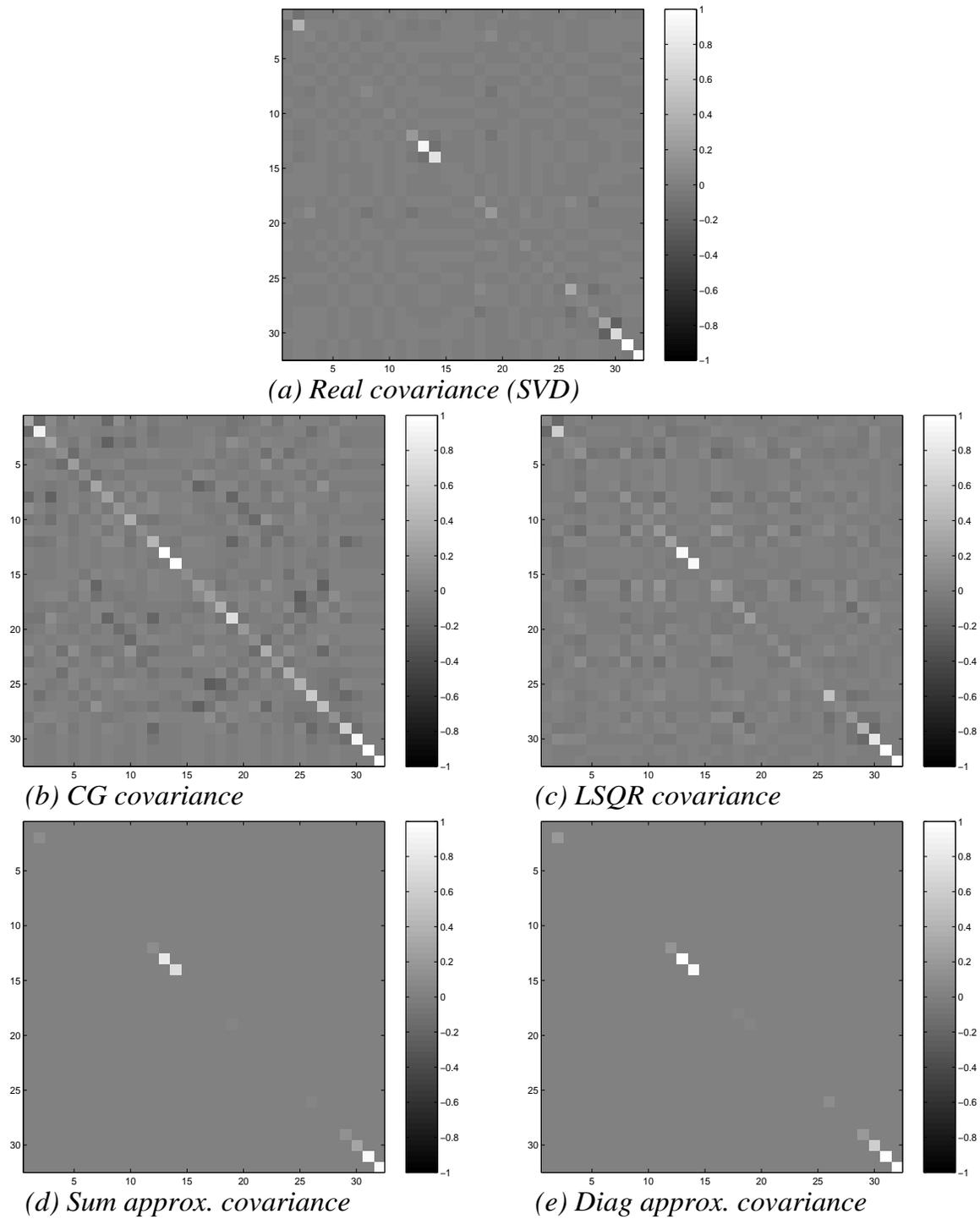


Figure 9: (a) Real covariance matrix calculated by SVD; (b)(c) Covariance matrices approximated by iterative methods; (d)(e) Covariance matrices approximated by model space weight methods. `barbara-covariance` [NR]

- Berkhout, A. J., 1997b, Pushing the limits of seismic imaging, part II: Integration of prestack migration, velocity estimation, and avo analysis: *Geophysics*, **62**, no. 3, 954–969.
- Berryman, J. G., 2001a, Analysis of approximate inverses in tomography I. Resolution analysis of common inverses: *Optimization and Engineering*, **1**, no. 1, 87–115.
- Berryman, J. G., 2001b, Analysis of approximate inverses in tomography II. Iterative inverses: *Optimization and Engineering*, **1**, no. 1, 437–473.
- Biondi, B., 1998, 3-D seismic imaging: *SEP-98*, 1–204.
- Bolte, J. F. B., and Verschuur, D. J., 1998, Aspects of focusing operator updating: 68th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 1604–1607.
- Claerbout, J., and Nichols, D., 1994, Spectral preconditioning: *SEP-82*, 183–186.
- Hegge, R. F., 2000, Seismic macromodel estimation by inversion of focusing operators: Ph.D. thesis, Delft University of Technology.
- Hestenes, M. R., and Stiefel, E., 1952, Methods of conjugate gradients for solving linear systems: *J. Res. Nat. Bur. Stan.*, **B**, no. 49, 409–436.
- Kabir, M. M. N., and Verschuur, D. J., 2000, A constrained parametric inversion for velocity analysis based on CFP technology: *Geophysics*, **65**, no. 4, 1210–1222.
- Menke, W., 1984, *Geophysical Data Analysis: Discrete Inverse Theory*: Academic Press, Orlando.
- Paige, C. C., and Saunders, M. A., 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: *ACM Transactions on Mathematical Software*, **8**, no. 1, 43–71.
- Thorbecke, J. W., 1997, *Common Focus Point Technology*: Ph.D. thesis, Delft University of Technology.
- Yao, Z. S., Roberts, R. G., and Tryggvason, A., 1999, Calculating resolution and covariance matrices for seismic tomography with the LSQR method: *Geophysical Journal International*, **138**, 886–894.

Born-compliant image perturbation for wave-equation migration velocity analysis

Paul Sava and Biondo Biondi¹

ABSTRACT

Wave-equation migration velocity analysis produces wrong results if it starts from an image perturbation which is not compliant with the assumed Born approximation. Earlier attempts to correct this problem lead to either unreliable or hard to implement solutions. In this paper, we present a new method designed to construct image perturbations that are always compliant with the Born approximation. This new method is robust, easy to implement, and produces results that are consistent with those obtained using the ideal operators.

INTRODUCTION

The fundamental idea of wave-equation migration velocity analysis (WEMVA) is that we can establish a *linear* relation between a small perturbation in the slowness field and the corresponding perturbation in the image. Therefore, given an image perturbation, we can invert for the slowness perturbation.

The main challenge of WEMVA is to construct a correct image perturbation that can be inverted for slowness. This image need not be fully accurate, but ought to provide the correct *direction* and *magnitude* of the change.

In our early tests (Sava and Biondi, 2000), we construct the image perturbation using Prestack Stolt Residual Migration (PSRM) (Stolt, 1996; Sava, 2000). In summary, this residual migration method provides updated images for new velocity maps that correspond to a fixed ratio of the new velocity with respect to the original velocity map.

The main disadvantage of building the image perturbation using PSRM is that, if the velocity ratio parameter (γ) is too large, there is a good chance for the reference and the updated images to get out of phase. In other words, a large change in velocity violates the Born approximation. The end result is that the image perturbation computed by the forward operator and the one computed by residual migration are fundamentally different, and can have contradictory behaviors when using the Born WEMVA operator for inversion.

In previous reports, we have presented various attempts to solve this problem. All these attempts were related one way or another to the idea of scaling down the change in the γ

¹email: paul@sep.stanford.edu, biondo@sep.stanford.edu

map, thus obeying the Born approximation, followed by scaling up of the inverted slowness map. Although these approaches were successful in several examples, none of them addressed the fundamental problem: does the (scaled) image perturbation created by residual migration match the one obtained by the forward WEMVA operator? Furthermore, scaling and rescaling cannot lead to a robust method, since they involve ad hoc processes and since the inversion problem we are trying to address is already highly nonlinear.

In this paper, we present a new method that can be used to create image perturbations for WEMVA. The two main goals here are

- to create an image perturbation that is compatible with the one computed using the forward WEMVA operator, and
- to create the image perturbation directly from the background image, and therefore compliant with the Born approximation.

We begin with a discussion of the WEMVA scattering operator, and continue with the derivation of our new method. Next we show a complex synthetic example and provide a discussion of the future work and of the problems that are still unsolved.

IMAGE PERTURBATION BY WEMVA

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

$$\mathcal{W}_{z+\Delta z} = \mathcal{W}_z e^{-ik_z \Delta z}, \quad (1)$$

where the depth wavenumber k_z depends linearly through a Taylor series expansion on its value in the reference medium (k_z^o) and the slowness in the depth interval from z to $z + \Delta z$, $s_o(z)$ and $s(x, y, z)$:

$$k_z \approx k_z^o + \left. \frac{dk_z}{ds} \right|_{s=s_o} \Delta s, \quad (2)$$

where, by definition, $\Delta s = s - s_o$.

If we denote the wavefield downward continued through the reference velocity as $\mathcal{W}_{z+\Delta z}^o$

$$\mathcal{W}_{z+\Delta z}^o = \mathcal{W}_z^o e^{-ik_z^o \Delta z}, \quad (3)$$

we obtain

$$\mathcal{W}_{z+\Delta z} = \mathcal{W}_{z+\Delta z}^o e^{-i \left. \frac{dk_z}{ds} \right|_{s=s_o} \Delta s \Delta z}. \quad (4)$$

The Born approximation linearizes the phase-shift exponential ($e^x \approx 1 + x$), such that we can write

$$\mathcal{W}_{z+\Delta z} \approx \mathcal{W}_{z+\Delta z}^o - i \mathcal{W}_{z+\Delta z}^o \left. \frac{dk_z}{ds} \right|_{s=s_o} \Delta s \Delta z. \quad (5)$$

Therefore, at any particular depth level, the wavefield perturbation $\Delta \mathcal{W}$ is

$$\Delta \mathcal{W} \approx -i \Delta z \mathcal{W} \left. \frac{dk_z}{ds} \right|_{s=s_o} \Delta s, \quad (6)$$

which we can also write as

$$\Delta \mathcal{W} \approx \frac{d\mathcal{W}}{dk_z} \left. \frac{dk_z}{ds} \right|_{s=s_o} \Delta s. \quad (7)$$

The image perturbation is simply obtained from the wavefield perturbation by summation over frequencies:

$$\Delta \mathcal{R} = \sum_{\omega} \Delta \mathcal{W}. \quad (8)$$

IMAGE PERTURBATION BY RESIDUAL MIGRATION

Residual migration can also be used to create an image perturbation. In its simplest form, we can build it as a difference between an *improved* image (\mathcal{R}) and the *reference* image (\mathcal{R}_o)

$$\Delta \mathcal{R} = \mathcal{R} - \mathcal{R}_o, \quad (9)$$

where \mathcal{R} is derived from \mathcal{R}_o as a function of the parameter γ , which is the ratio of the original and improved velocities (Sava, 2000). Of course, the improved velocity map is unknown explicitly, but it is described indirectly by the ratio map of the two velocities.

If we define $\Delta \gamma = \gamma - 1$, we can also write the discrete version of the image perturbation as

$$\Delta \mathcal{R} \approx \frac{\mathcal{R} - \mathcal{R}_o}{\gamma - 1} \Delta \gamma, \quad (10)$$

equation which can be written in differential form as

$$\Delta \mathcal{R} \approx \left. \frac{d\mathcal{R}}{d\gamma} \right|_{\gamma=1} \Delta \gamma, \quad (11)$$

or, equivalently, using the chain rule, as

$$\Delta \mathcal{R} \approx \frac{d\mathcal{R}}{dk_z} \left. \frac{dk_z}{d\gamma} \right|_{\gamma=1} \Delta \gamma, \quad (12)$$

where k_z is the depth wavenumber defined for PSRM.

The Equations (7) and (12) are very similar, which comes at no surprise since they effectively represent the same thing: the perturbation of the image given a perturbation of the slowness field, or equivalently, a perturbation of the γ (ratio) field. We will use Equation (12)

to create the image perturbation, which we will then backproject in the slowness space using the adjoint of the WEMVA equation (7).

Equation (12) offers the possibility to build the image perturbation directly. We achieve this by computing three elements: the derivative of the image with respect to the depth wavenumber, and two weighting functions, one for the derivative of the depth wavenumber with respect to the velocity ratio parameter (γ), and the other one for the magnitude of the $\Delta\gamma$ perturbation from the reference to the improved image.

Firstly, the image derivative in the Fourier domain, $\frac{d\mathcal{R}}{dk_z}$, is straightforward to compute in the space domain as

$$\left. \frac{d\mathcal{R}}{dk_z} \right|_{\gamma=1} = -iz\mathcal{R}_o. \quad (13)$$

The derivative image is nothing but the imaginary part of the migrated image, scaled by depth.

Secondly, we can obtain the weighting representing the derivative of the depth wavenumber with respect to the velocity ratio parameter, $\left. \frac{dk_z}{d\gamma} \right|_{\gamma=1}$, starting from the double square root (DSR) equation written for prestack Stolt residual migration (Sava, 2000):

$$\begin{aligned} k_z &= k_z^s + k_z^r \\ &= \frac{1}{2}\sqrt{\gamma^2\mu^2 - |\mathbf{k}_s|^2} + \frac{1}{2}\sqrt{\gamma^2\mu^2 - |\mathbf{k}_r|^2}, \end{aligned}$$

where μ is given by the expression:

$$\mu^2 = \frac{\left[4(k_z^o)^2 + (|\mathbf{k}_r| - |\mathbf{k}_s|)^2\right] \left[4(k_z^o)^2 + (|\mathbf{k}_r| + |\mathbf{k}_s|)^2\right]}{16(k_z^o)^2}. \quad (14)$$

The derivative of k_z with respect to γ is

$$\frac{dk_z}{d\gamma} = \gamma \left(\frac{\mu^2}{4k_z^s} + \frac{\mu^2}{4k_z^r} \right), \quad (15)$$

therefore

$$\left. \frac{dk_z}{d\gamma} \right|_{\gamma=1} = \frac{\mu^2}{2\sqrt{\mu^2 - |\mathbf{k}_s|^2}} + \frac{\mu^2}{2\sqrt{\mu^2 - |\mathbf{k}_r|^2}}. \quad (16)$$

Finally, $\Delta\gamma$ can be picked from the set of residually migrated images at various values of the parameter γ (Sava, 2000). The main criterion that should be used is the flatness of the angle-domain image gathers, although in principle other derived parameters, such as stack power or semblance, can be used as well.

EXAMPLES

In this section, we demonstrate the method on a synthetic example. The reflectivity model (Figure 1) consists of two flat reflectors surrounding a set of reflectors dipping at 45 degrees.

Figure 1: Reflectivity model.
paul1-rr [ER]

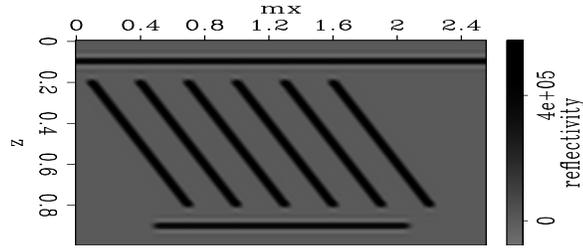
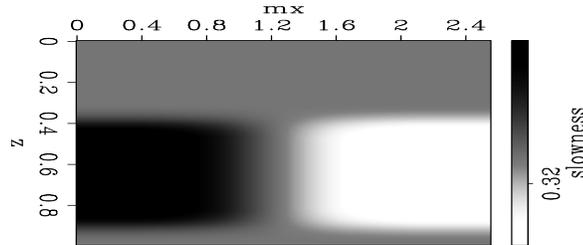
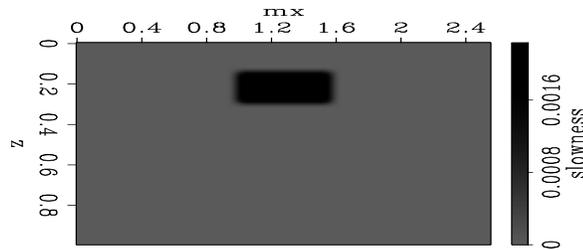


Figure 2: Background slowness model.
paul1-s1 [ER]



The background velocity (Figure 2) is characterized by strong lateral variation to provide a somewhat complex background wavefield. The image perturbation (Figure 3) consists of a rectangular block in the upper part of the section, which creates perturbations in the image both on the flat and on the dipping reflectors. We change the magnitude of this anomaly from small values that satisfy the Born approximation, to high values that clearly violate the linear assumptions.

Figure 3: Slowness perturbation.
paul1-ds [ER]



We use the true slowness to model the data, and then we use the background slowness to obtain the reference image and the reference wavefield. From the slowness perturbation, we use the forward WEMVA operator to create the *ideal* image perturbation (Figure 4). Backprojection by the adjoint WEMVA operator gives the slowness update in Figure 5, and inversion gives the slowness in Figure 6. Figures 4 and 5 represent the benchmark against which we want to compare our new image perturbation method.

Figure 7 represents the image perturbation created using Equation (12) where we consider that $\Delta\gamma = 1$. Of course, this is not our image perturbation, but it only gives us a rough idea of how we should pick the actual perturbation from after weighting with $\Delta\gamma$. As expected, the magnitude of this component of the image increases with depth.

Next we multiply the image in Figure 7 with the $\Delta\gamma$ weight (Figure 8) and obtain an image perturbation (Figure 9) which is comparable in shape and magnitude with the ideal perturbation (Figure 4). Backprojection by the adjoint WEMVA operator gives the slowness update in Figure 10, and inversion gives the slowness in Figure 11.

Figure 4: Ideal image perturbation obtained by the forward WEMVA operator. [paul1-di.wei.01](#) [ER]

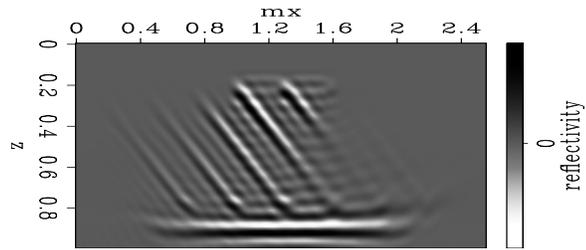


Figure 5: Slowness backprojection obtained from the ideal image perturbation using the adjoint WEMVA operator. [paul1-bds.wei.01](#) [ER]

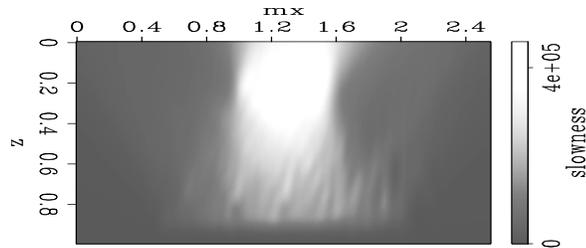


Figure 6: Slowness inversion obtained from the ideal image perturbation using the WEMVA operator. [paul1-ids.wei.01](#) [ER]

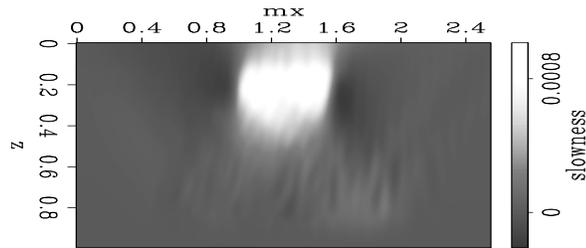


Figure 7: Image perturbation produced by Equation (12). The magnitude of the perturbation increases with depth. [paul1-di.ana.full](#) [ER]

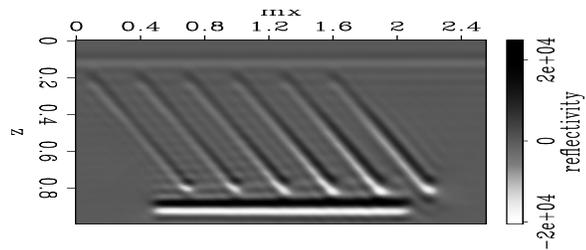


Figure 8: $\Delta\gamma$ in Equation (12). [paul1-mask](#) [ER]

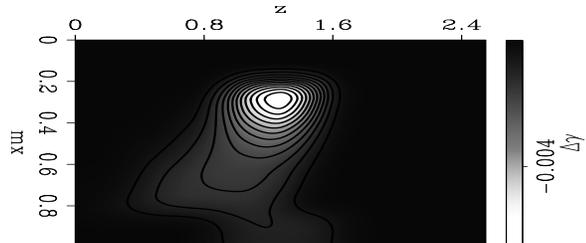


Figure 9: Analytical image perturbation obtained by multiplication of the image in Figure 7 with the weight in Figure 8. [paul1-di.ana.01](#) [ER]

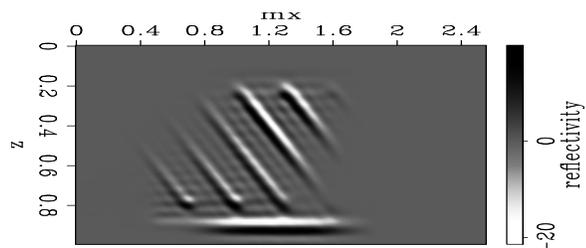


Figure 10: Slowness backprojection obtained from the analytical image perturbation using the adjoint WEMVA operator. `paul1-bds.ana.01` [ER]

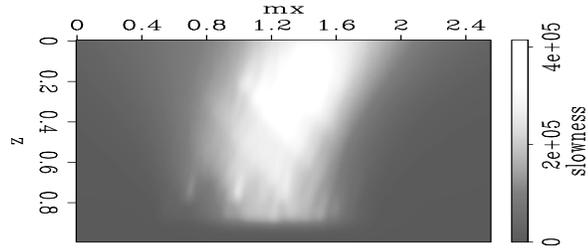
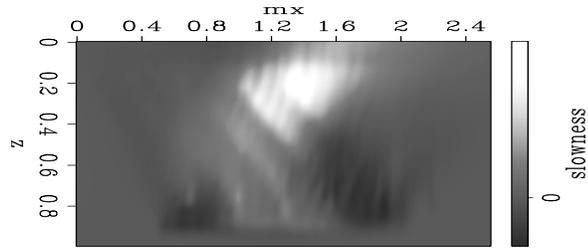


Figure 11: Slowness inversion obtained from the analytical image perturbation using the WEMVA operator. `paul1-ids.ana.01` [ER]



Finally, Figures 12, 13, and 14 are comparisons of the various methods used to compute the image perturbation, and the corresponding slowness model. All figures are presented in landscape mode and have three columns and four rows.

- The first column contains the real part of the image perturbation ($\Delta\mathcal{R}$), the second column contains the slowness model (Δs) obtained using backprojection of the image perturbation, and the third column contains the inverted slowness model obtained from the same image perturbation.
- The first row represents $\Delta\mathcal{R}$ and Δs computed using the WEMVA forward and adjoint operators, the second row represents $\Delta\mathcal{R}$ and Δs in the case in which we narrow down the angular aperture in the image perturbation, the third row represents $\Delta\mathcal{R}$ and Δs of our new approach, and the fourth row represents $\Delta\mathcal{R}$ and Δs obtained using the PSRM method.

Each of the Figures 12, 13, and 14 corresponds to a different magnitude of the slowness anomaly, everything else remaining unchanged. The main comparison is between rows 2 and 3 of the composite image. We can observe that as we increase the magnitude of the anomaly, the shapes of $\Delta\mathcal{R}$ and Δs remain roughly unchanged, although the magnitudes of both the image and slowness perturbations increase. Since, by construction, we obey the Born approximation, we do not observe any out-of-phase effects.

In contrast, when we compare rows 3 and 4, that is our new method and the PSRM approach, we observe that, for the small anomalies, the image and slowness perturbations are very similar, but at high values of the anomaly, the PSRM $\Delta\mathcal{R}$ gets out of phase with respect to the background wavefield, and therefore the backprojected or inverted Δs either cancels or changes sign altogether. This is exactly the kind of effect we are trying to avoid using our new approach.

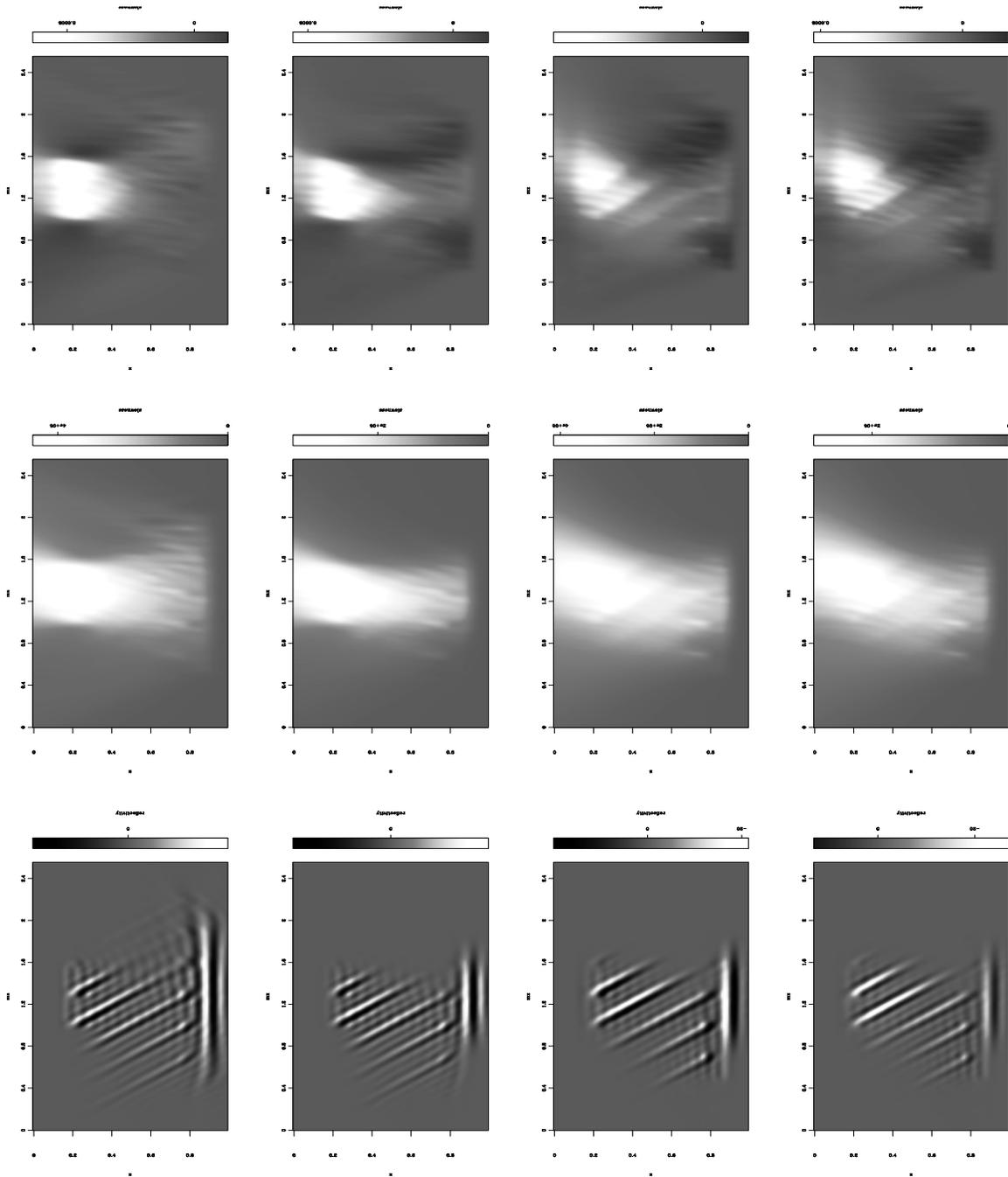


Figure 12: View as 4 rows and 3 columns. The left column represents the image perturbations, the middle represents backprojection using the WEMVA operator, and the right column represents unconstrained inversion in 3 iterations using the WEMVA operator. The top row is using the ideal image perturbation, the second row is using the narrow aperture image perturbation, the third row is using our new approach, and the fourth row the PSRM approach.

paul1-block.01 [CR]

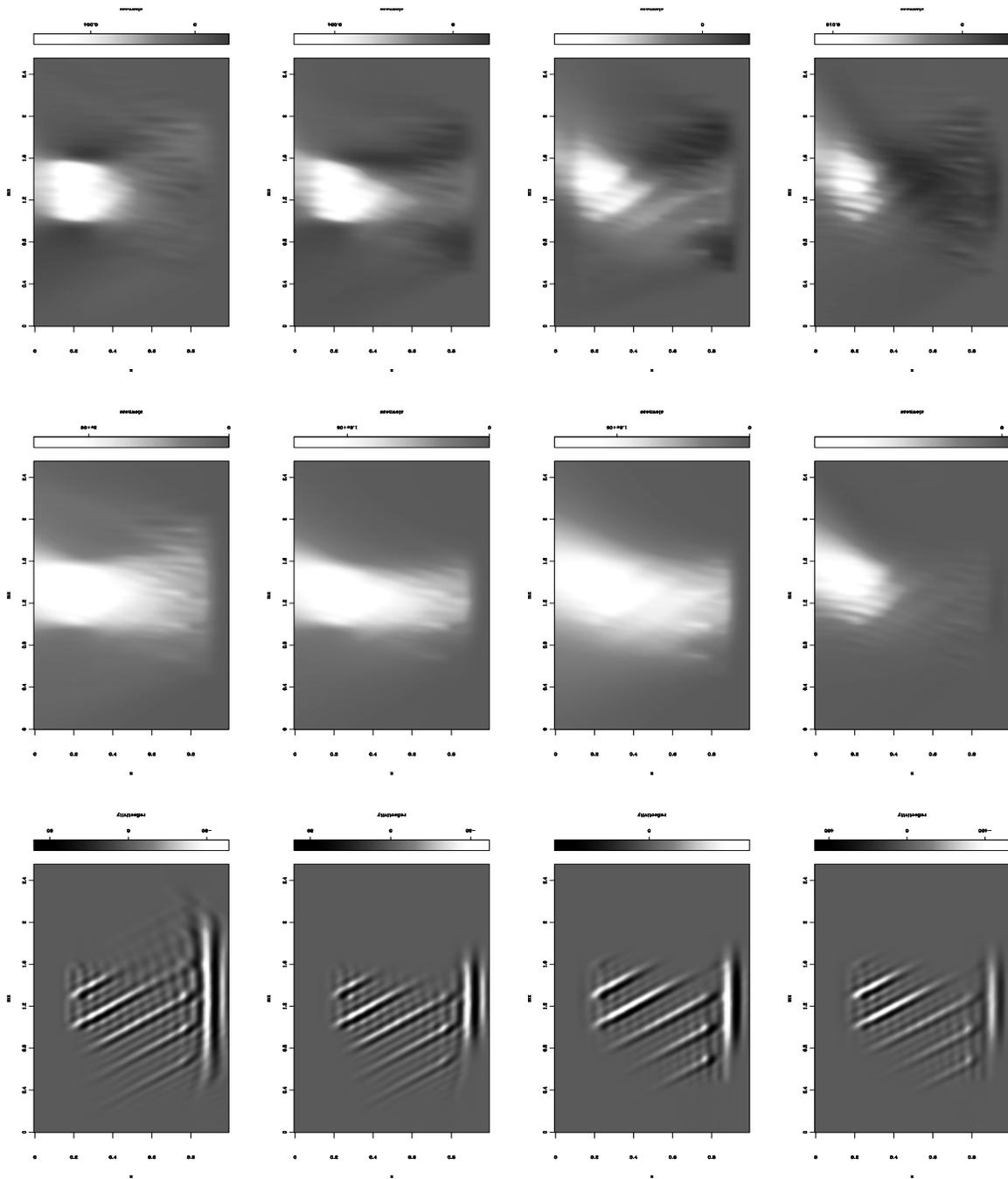


Figure 13: View as 4 rows and 3 columns. The left column represents the image perturbations, the middle represents backprojection using the WEMVA operator, and the right column represents unconstrained inversion in 3 iterations using the WEMVA operator. The top row is using the ideal image perturbation, the second row is using the narrow aperture image perturbation, the third row is using our new approach, and the fourth row the PSRM approach.

paul1-block.06 [CR]

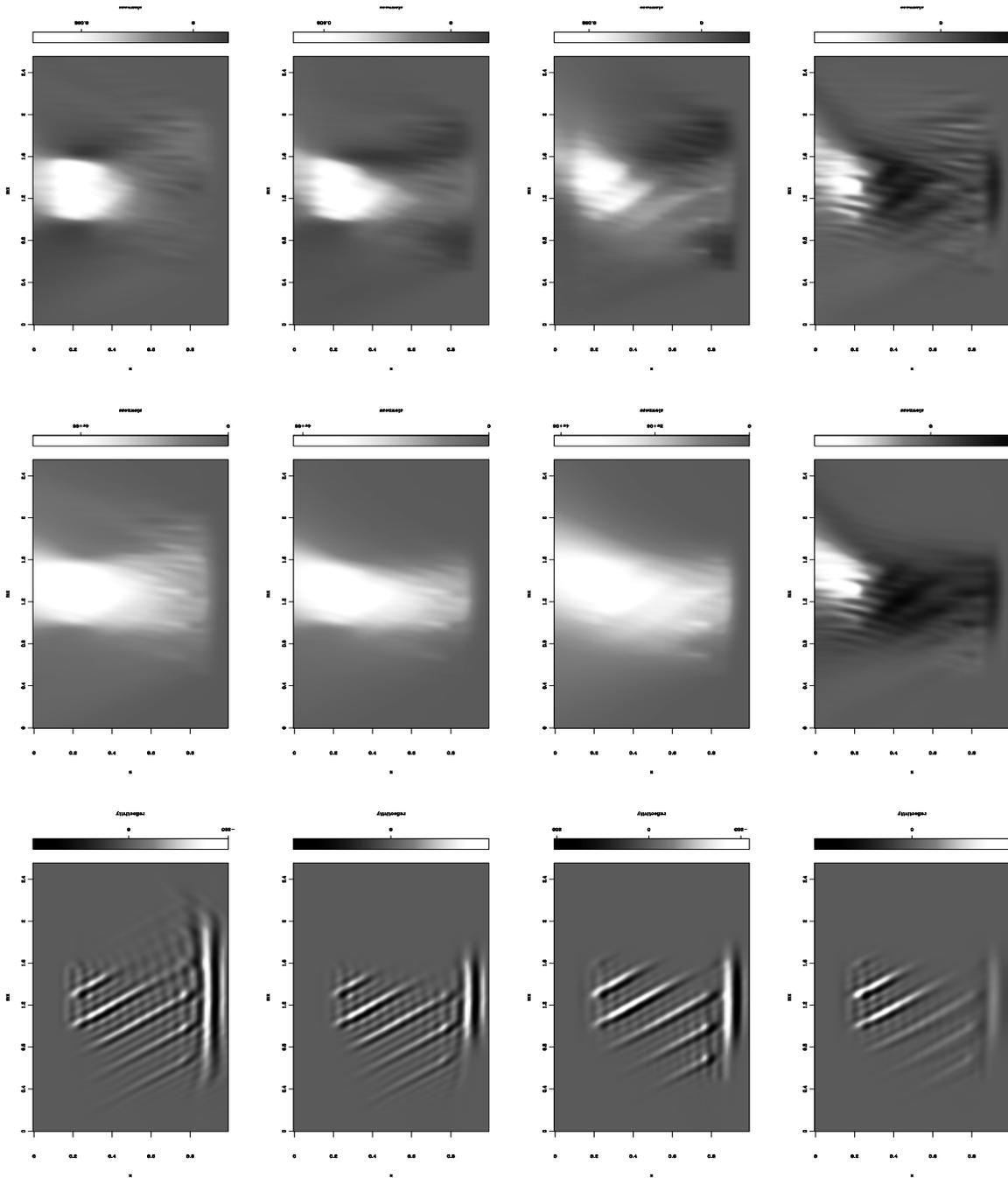


Figure 14: View as 4 rows and 3 columns. The left column represents the image perturbations, the middle represents backprojection using the WEMVA operator, and the right column represents unconstrained inversion in 3 iterations using the WEMVA operator. The top row is using the ideal image perturbation, the second row is using the narrow aperture image perturbation, the third row is using our new approach, and the fourth row the PSRM approach.

`paul1-block.11` [CR]

FUTURE WORK

At this moment, the main problem that is still not fully solved is that of picking the $\Delta\gamma$ (Figure 8) distribution. For the examples in this paper, we have artificially constructed this weight from information provided by the forward WEMVA operator, information which we would not normally have when dealing with real data.

In theory, this weight should be consistent with the flatness of the image gathers: smaller values for nearly flat events in the image gathers, and larger ones for events with higher moveout. As we mentioned earlier, we could use other parameters related to this moveout, such as stack power or semblance. In practice, however, any one of these approaches requires picking, which is a very tedious task, especially in 3-D. Perhaps, some kind of automatic picking (Clapp, 2001) is the answer, but this issue remains to be solved in future work.

CONCLUSIONS

We have presented a new approach to the construction of image perturbations for velocity analysis using WEMVA. This method directly constructs the image perturbation from the background image, and is always compliant with the Born approximation which is the underlying assumption of WEMVA. We show that, given correct scaling, we can obtain slowness anomalies that are fully consistent with those obtained by the application of the forward and adjoint WEMVA operators.

REFERENCES

- Clapp, R., 2001, Ray-based tomography with limited picking: SEP-**110**, 103–112.
- Sava, P., and Biondi, B., 2000, Wave-equation migration velocity analysis: Episode II: SEP-**103**, 19–47.
- Sava, P., 2000, Prestack Stolt residual migration for migration velocity analysis: 70th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 992–995.
- Stolt, R. H., 1996, Short note—a prestack residual time migration operator: Geophysics, **61**, no. 02, 605–607.

Short Note

Ray-based tomography with limited picking

Robert G. Clapp¹

INTRODUCTION

In ray-based reflection tomography picking reflectors is an integral and painful part of the process (Clapp, 2001; van Trier, 1990; Stork, 1992; Kosloff et al., 1996). The general 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, which in turn leads to back projecting information to the wrong portion of the model space, seriously hampering the inversion.

In this paper I present a method to eliminate, or at least significantly reduce, the need for reflector picking. I calculate a dip field and coherency from a migrated image by first using the plane-wave estimator from Claerbout (1992) and later used by Bednar (1997). I next refine the dip estimate using the methodology described in Fomel (2000). I then automatically select back projection points based on dip coherency and semblance strength. The method is applied to a 2-D North Sea Dataset. The turn around time is reduced significantly and the overall image quality is equal to or better than conventional reflector picking approach.

THEORY

Figure 1 shows the typical flow for ray-based reflection tomography.

- First a set of reflectors are picked. Then residual moveout analysis is performed along the reflectors.
- This moveout is characterized by one or more parameters.
- The reflectors are also used as the basis for back projection points. The normal of the picked reflectors is used either for calculating the initial takeoff angle (if shooting up

¹email: bob@sep.stanford.edu

from the reflector) or for ensuring Snell's law is obeyed at the reflector (if shooting down to the reflector).

- Finally, time errors are estimated for each ray pair (Stork, 1992).

Two of the largest potential sources of error in this estimation scheme are an inaccurate dip estimate (causing information to back projected into the wrong portion of model space) and inaccurate description of moveout. The first type of error is the result of a poor description of the reflectors, which may be the result of overuse of an auto-picker or too little or too much smoothing of reflector positions. Poor moveout description is often the result of extending reflectors into areas with low signal-noise ratio where moveout analysis gives unreliable information. Both problems can be attenuating with significant human QCing, but will increase turnaround time.

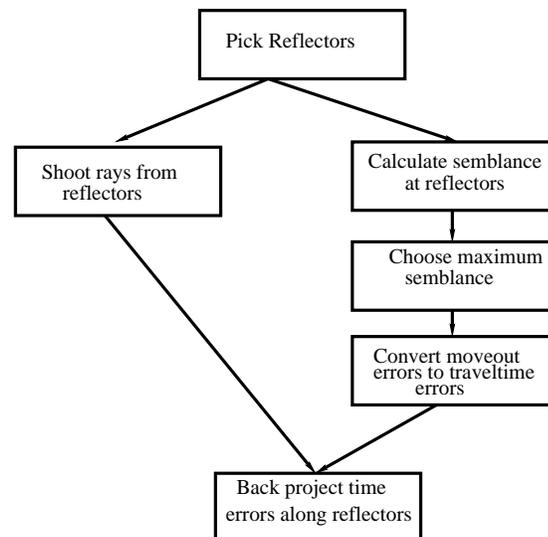


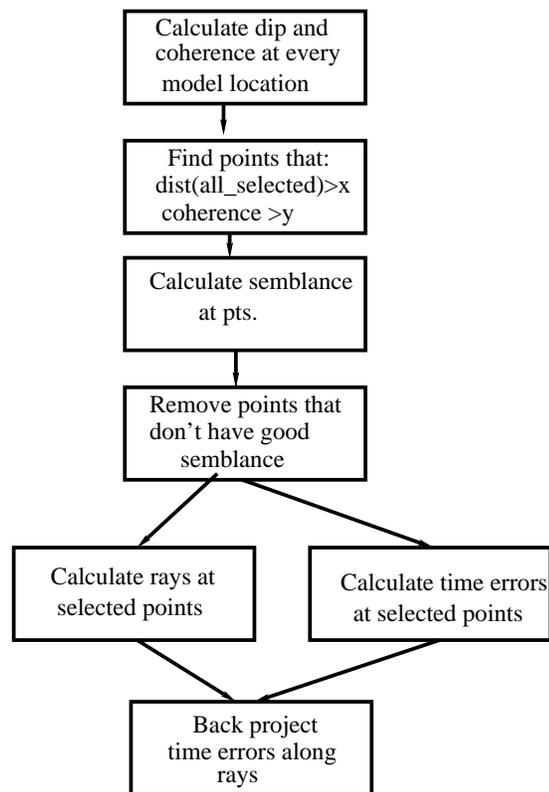
Figure 1: A typical ray-based reflection tomography loop. [bob1-old]
[NR]

Figure 2 shows an alternate ray-based tomography flow. Rather than using picked reflectors as the basis for back projection locations, points are selected according to reliability factors. First dip and coherency of the migrated image is calculated at each image location. For an initial dip and coherence estimate I take a window around each model location. I calculate the best single dip within the region, and the coherence of that dip, using the method described in Claerbout (1992). I then use this as an initial dip estimate for the non-linear, space varying dip estimation procedure described in Fomel (2000).

Likely back projection points are then automatically selected by finding model locations that meet some specified dip coherence, amplitude, and distance from other selected points. To get the 'best' points in each region, these criteria are slowly relaxed (e.g. the first pass might look for points above the 90th percentile in amplitude and dip coherence, while the last pass might drop both these criteria to the 50th percentile.).

At each initially selected point semblance analysis is performed. Points that don't have good semblance (large semblance value and a definite maximum) are discarded. The remaining points are then used.

Figure 2: Back-projection scheme used in this paper. Note the absence of reflector picking. `bob1-new` [NR]



DATA EXAMPLE

To test the methodology I applied it on a 2-D North Sea line taken from a 3-D volume, the same dataset used in (Clapp, 2000, 2001). There is a significant 3-D component to the data, especially from the salt structure. Figure 3 shows the initial velocity model. The initial velocity model was created by smoothing to an extreme the S.M.A.R.T² method (Jacobs et al., 1992; Ehinger and Lailly, 1995). Figure 4 shows the initial migration (by doing split-step downward continuation) of the data and Figure 5 show every 10th CRP gather. Note how there is significant residual moveout throughout the model.

From this initial migrated image I calculated the dip (Figure 6) and coherence (Figure 7). To calculate the initial back projection points I selected points above 30th percentile in image amplitude, 45th percentile in dip coherence, and at least four model points away from all other selected points. I then calculated semblance at each point and parsed from the list points with semblance below .16 and whose maximum was less than 40% below the average semblance. Figure 8 shows the final location of back projection points. Note how there are few points near the salt body and in the lower portion of the image.

As image quality improves, the automatic selector slowly uses more and more model

²Sequential Migration-Aided Reflection Tomography - KIM (Kinematic Inversion Methods), IFP consortium

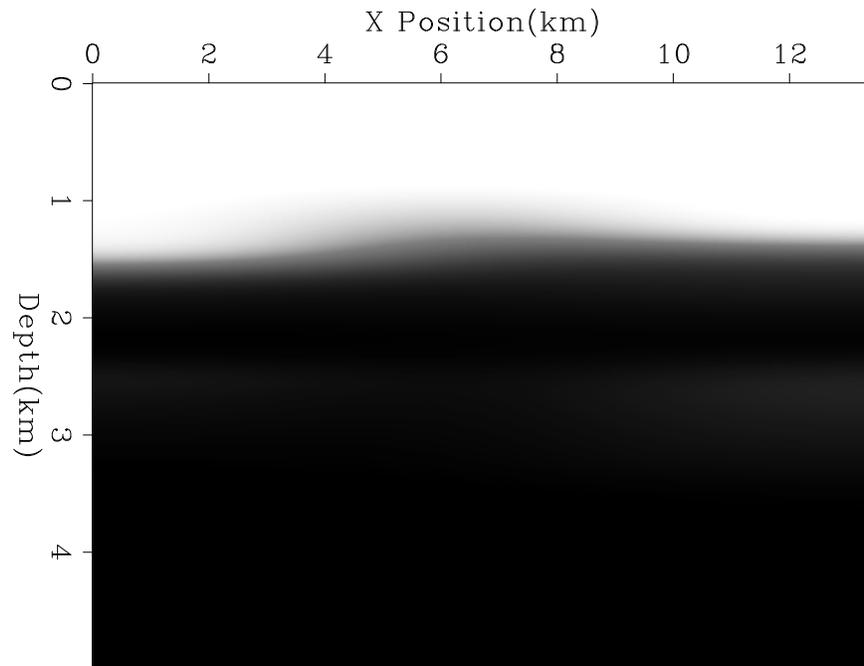


Figure 3: Initial velocity model. `bob1-vel0` [CR]

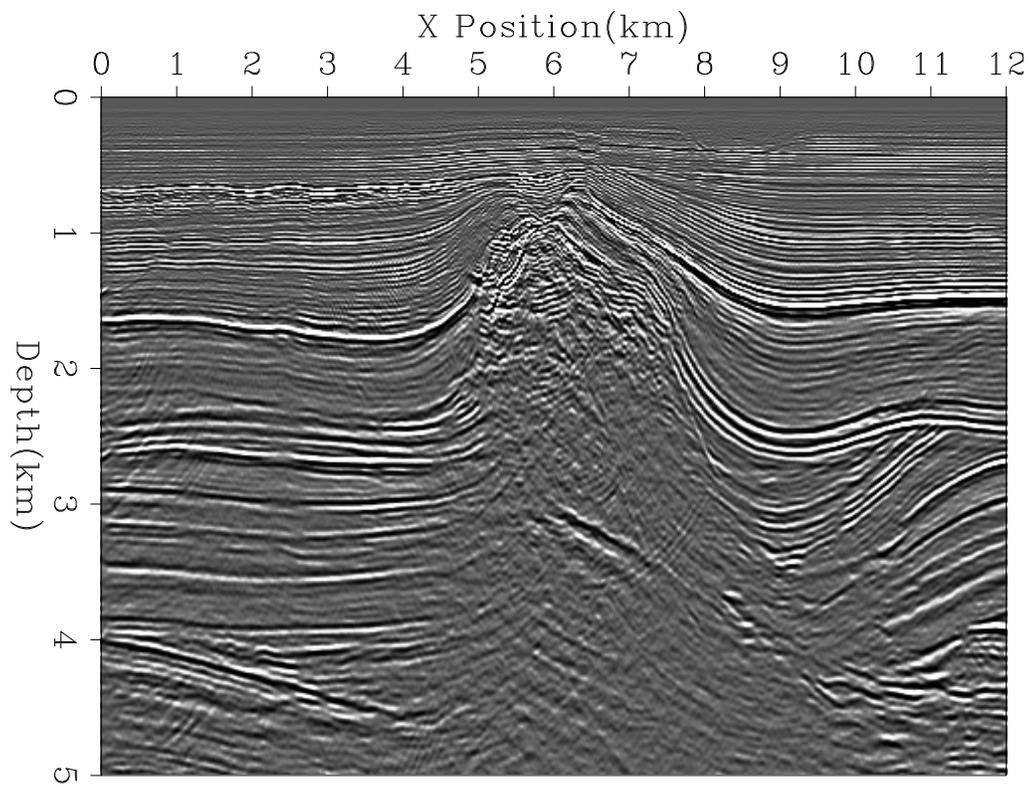


Figure 4: Initial migration using the velocity model shown in Figure 3. `bob1-image.vel0` [CR]

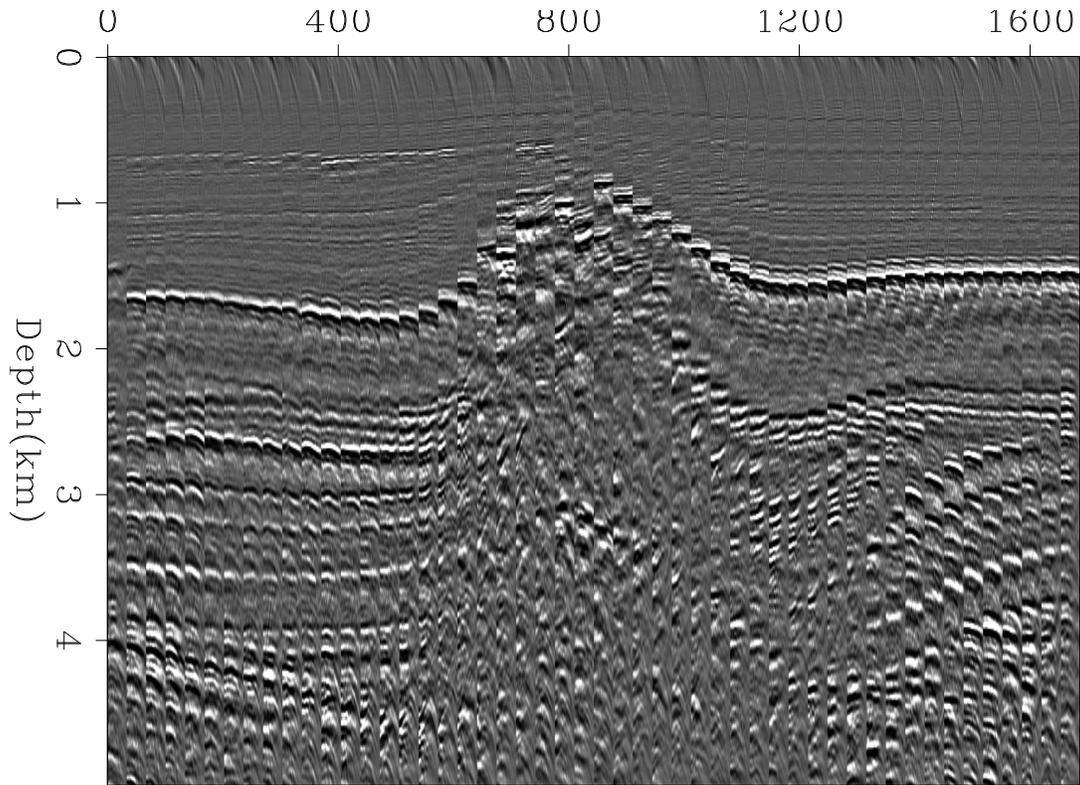


Figure 5: Every 10th CRP gather of the initial migration (Figure 4) using the velocity model shown in Figure 3. `bob1-mig.vel0` [CR]

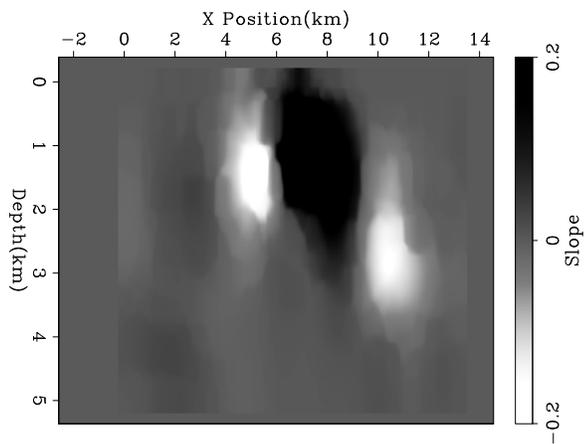


Figure 6: The calculated dip field for the initial migration image. `bob1-dipxz.vel0` [CR]

Figure 7: The calculated dip coherence for the initial migration image.
`bob1-coher.vel0` [CR]

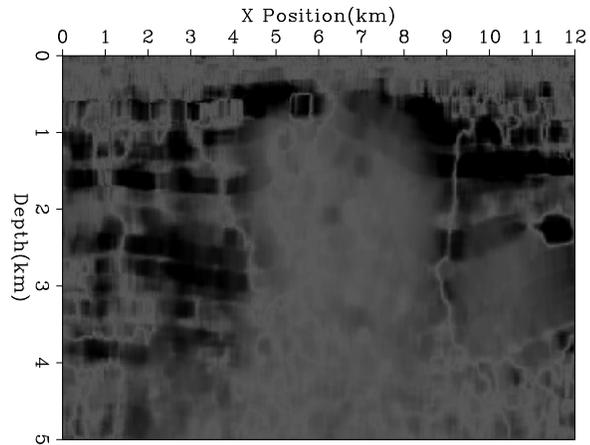
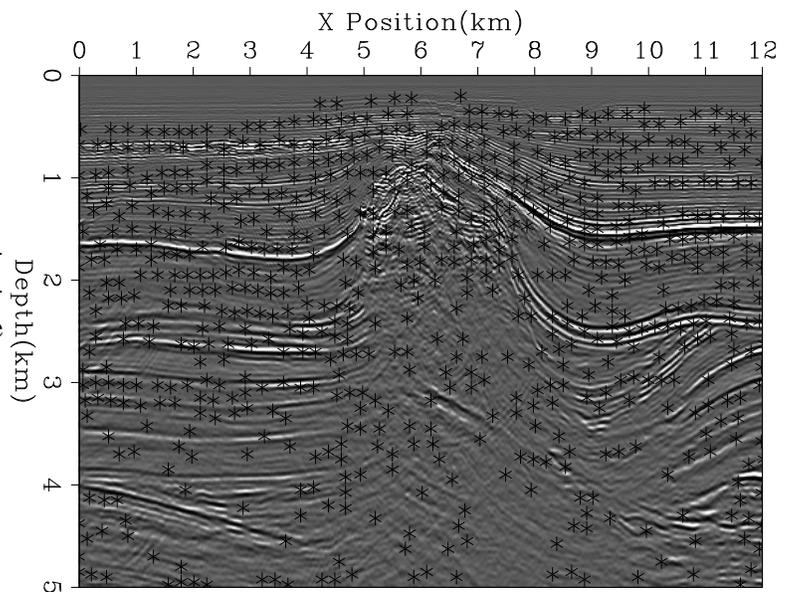


Figure 8: The selected back projection points. Note how the points generally follow reflectors and avoid the salt structure. `bob1-pts.vel0` [CR]



points. Figure 9 shows the points selected after the five iteration of tomography. Note how there are now many more points in the lower portion of the model and closer to the salt. The velocity model after five iterations can be seen in Figure 10 and the resulting migrated image in Figure 11. Figure 12 shows the CRP gathers of the fifth migration. Note how they are significantly flatter than in Figure 12.

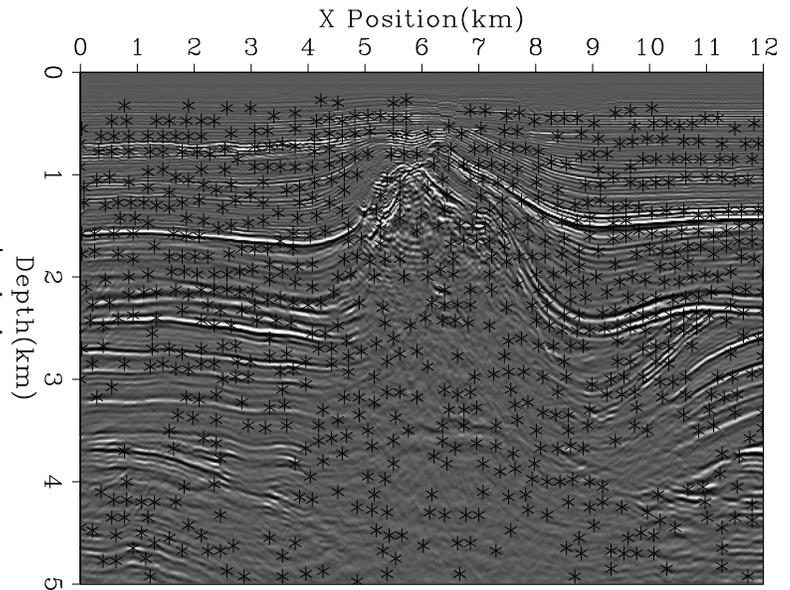


Figure 9: The selected back projection points after five iterations. Note the difference from Figure 8.

bob1-pts.final [CR]

CONCLUSIONS

Reflector picking is one of the main bottlenecks in ray-based tomography. By automatically selecting back projection points based on dip coherency and high semblance, picking can be avoided. The results on a 2-D dataset are encouraging.

REFERENCES

- Bednar, J. B., 1997, Least squares dip and coherency attributes: SEP-95, 219–225.
- Claerbout, J. F., 1992, Earth Soundings Analysis: Processing versus Inversion: Blackwell Scientific Publications.
- Clapp, R., 2000, Tau tomography with steering filters: 3-D field data example, preliminary result: SEP-105, 123–138.
- Clapp, R. G., 2001, Geologically constrained migration velocity analysis: Ph.D. thesis, Stanford University.
- Ehinger, A., and Lailly, P., 1995, Velocity model determination by the SMART method, part 1: Theory: 65th Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 739–742.

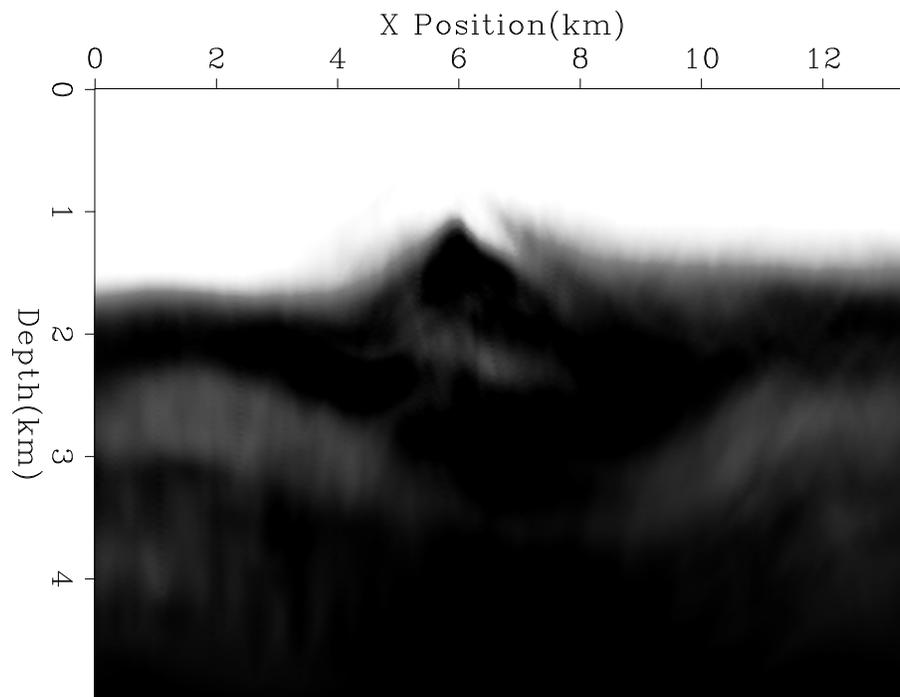


Figure 10: Final velocity model. `bob1-final` [CR]

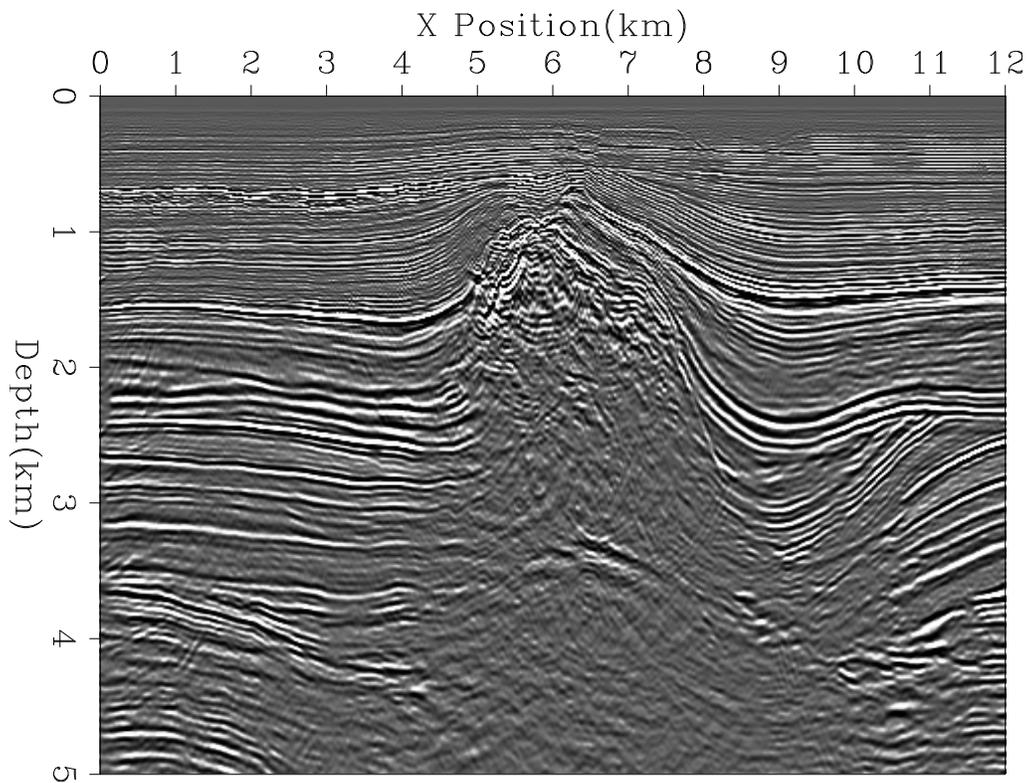


Figure 11: Final migration using the velocity model shown in Figure 10. `bob1-image.final` [CR]

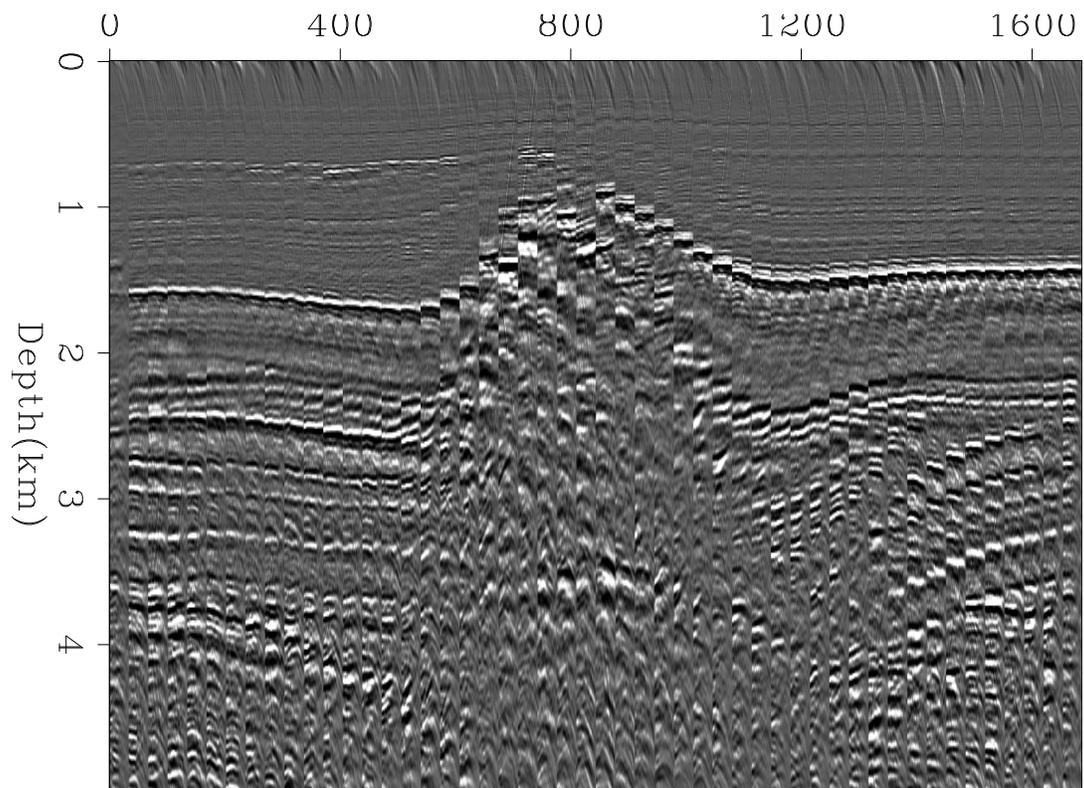


Figure 12: Every 10th CRP gather of the final migration (Figure 11) using the velocity model shown in Figure 10. `bob1-mig.final` [CR]

- Fomel, S., 2000, Applications of plane-wave destructor filters: *SEP-105*, 1–26.
- Jacobs, J. A. C., Delprat-Jannaud, F., Ehinger, A., and Lailly, P., 1992, Sequential migration-aided reflection tomography: A tool for imaging complex structures: 62nd Annual Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1054–1057.
- 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.
- 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.

A new multiscale prediction-error filter for sparse data interpolation

William Curry and Morgan Brown¹

ABSTRACT

Prediction-error filters (PEFs) have been used to successfully interpolate seismic data. Using conventional methods, PEFs often cannot be estimated on sparse, irregularly sampled data. We implement an algorithm in which we resample the data to various scales to estimate a single PEF. We show that compared to PEFs estimated from single data scales, our PEF provides a robust first guess for a nonlinear interpolation scheme.

INTRODUCTION

Data interpolation can be formulated as an inverse problem. If the data and model are in the same space, the data fitting goal is simply a mask operator, which guarantees that the model fits the data at the known points. The model is unconstrained where there are unknown data, leaving a family of possible models with a large nullspace. Unfortunately, there is an ambiguity in how to constrain the solution so that a reasonable result is produced. Claerbout (1999) suggests that the covariance of the known data can be used as the covariance of the model. The data covariance can be characterized by a prediction-error filter, which is estimated from the known data. In the second stage of this method, the PEF is then used to regularize the inversion and constrain the nullspace.

In some cases, the sparse data may be so sparsely sampled as to make conventional PEF estimation impossible. For the special case of interpolating between regularly sampled traces, Crawley (2000) spaces the coefficients of the PEF at multiple scales, and successfully interpolates aliased events. Fomel (2001) uses a nonlinear method to estimate dips within data. The method works well, but is computationally expensive. A test case for sparse data interpolation has been developed by Brown, et al. (2000), which consists of a single plane wave that is sparsely and irregularly sampled.

We develop a method that correctly interpolates a more difficult test case, and provides an overall strategy to interpolate sparse, irregular data when existing methods fail. To do this, we develop a PEF estimation scheme where a single PEF is estimated with multiple scales of regridded data, by simultaneously autoregressing for a common filter. We show that this method is more robust than estimating a PEF on a single scale of data, and provides more

¹email: bill@sep.stanford.edu, morgan@sep.stanford.edu

equations than estimating the PEF with multiple scales of filters. Once we estimate the PEF, we use the interpolated data as a starting guess for a nonlinear iterative method, which gives promising results. This method is then compared favorably to starting guesses based upon Laplacian interpolation and PEFs estimated from a single scale of data.

BACKGROUND

Sparse data interpolation is a very important problem in geophysics, due to the high cost associated with data collection. Irregular sampling introduces another level of complexity to the problem. Data interpolation can be implemented using a two stage linear method, the second of which is the minimization of the model when convolved with a filter. The first step is determining an appropriate filter, such as a prediction-error filter (PEF) (Claerbout, 1999).

A PEF can often be estimated by minimizing the output of convolution of an unknown filter with known data. The one dimensional case,

$$\mathbf{0} \approx \mathbf{r} = \begin{bmatrix} y_2 & y_1 & y_0 \\ y_3 & y_2 & y_1 \\ y_4 & y_3 & y_2 \\ y_5 & y_4 & y_3 \\ y_6 & y_5 & y_4 \end{bmatrix} \begin{bmatrix} 0 & \cdot & \cdot \\ \cdot & 1 & \cdot \\ \cdot & \cdot & 1 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix} + \begin{bmatrix} y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} \quad (1)$$

can easily be extended to multiple dimensions with the helical coordinate (Claerbout, 1998). Here a represents filter coefficients, y represents data, and \mathbf{r} is the residual, which we minimize. When there are unknown data, the equations containing missing data are not used. Very sparse data complicates the issue, since there are no fitting equations which determine a PEF. Looking at the above equation, if y_2 , y_3 and y_4 were missing, a PEF could not be determined.

A method to overcome the need for contiguous data is to stretch the filter (Crawley, 1998) so that the filter coefficients would fall upon data points. The stretching can be done at multiple scales, due the the scale-invariance of the PEF (Claerbout, 1999). This method works well as long as the data are regularly sampled, and the stretching is isotropic. If the data are irregularly sampled, the method fails, since the filter coefficients no longer fall upon data points, and again we are left without fitting equations.

$$\begin{matrix} \cdot & a & \cdot & b & \cdot & c & \cdot & d \\ \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \end{matrix}$$

A test case has been developed (Brown et al., 2000), which consists of a single plane wave oriented at 22.5° , irregularly sampled. We extend this test case, and add complexity to it by adding another plane wave with a different frequency and orientation. In addition we add a substantial amount of Gaussian noise. Crawley’s (1998) method of scaling the filter does not work for this case, as the irregular sampling leads to a lack of fitting equations. The sampling is much like the middle case in Figure 1 where the filter does not lie on known data, regardless of how much it is stretched.

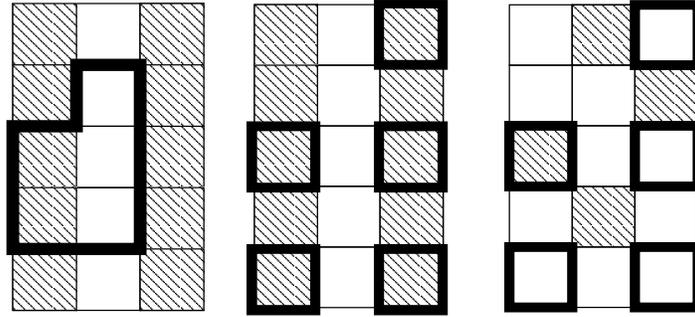


Figure 1: Figure of three cases for a regression equation. The diagonal lines represent gridded data, the white squares are empty bins, and the bold lines represent the PEF. Left: Original PEF on interlaced data, where an equation is not possible. Center: scale-expanded PEF on interlaced data, with a possible equation. Right: scale-expanded PEF on irregularly spaced data, no equation is possible. `bill1-peffail` [NR]

MULTISCALE PEF ESTIMATION

Instead of scaling the filter to fit the data, the data can be scaled so that it fits the filter. This is accomplished by a regridding algorithm, which we base upon normalized linear interpolation. The data can be regridded at multiple scales, so that the number of fitting equations increases. The success of this method is dependent upon the scale-invariance of the data instead of that of the PEF. If we assume stationarity, the regridding is an acceptable solution, until we get to very large bin sizes where adequate sampling becomes an issue.

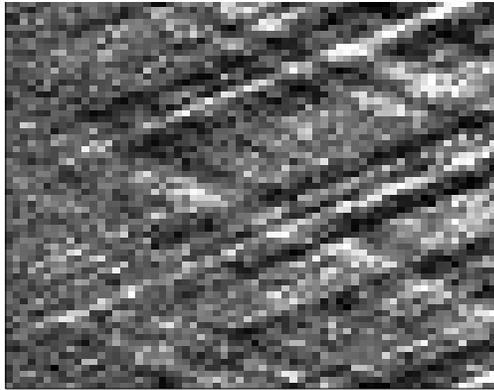
The calculation of a multiscale PEF can be described as multiple simultaneous PEF estimations, one for each scale, with its own mask,

$$\begin{bmatrix} \mathbf{K}_0\mathbf{D} \\ \mathbf{K}_1\mathbf{S}_1\mathbf{D} \\ \mathbf{K}_2\mathbf{S}_2\mathbf{D} \\ \dots \\ \mathbf{K}_n\mathbf{S}_n\mathbf{D} \end{bmatrix} \mathbf{f} = \begin{bmatrix} \mathbf{d} \\ \mathbf{S}_1\mathbf{d} \\ \mathbf{S}_2\mathbf{d} \\ \dots \\ \mathbf{S}_n\mathbf{d} \end{bmatrix}. \quad (2)$$

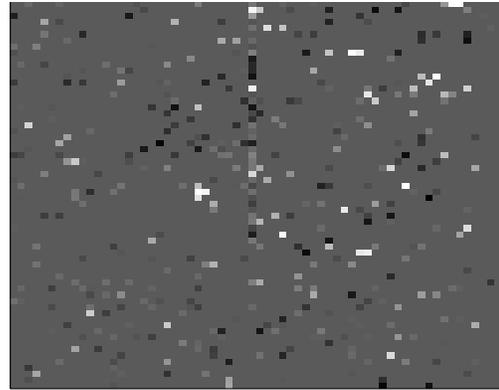
In equation (2), \mathbf{D} signifies convolution with the data, \mathbf{S}_n is a regridding matrix, which regrids to the n th scale. \mathbf{K}_n is a weighting vector which is 1 where all data are present in the equation at the n th scale and 0 where there are missing data, \mathbf{f} is the filter, and \mathbf{d} is the data.

There are multiple benefits to this approach of estimating the PEF. First, the data can be sampled very irregularly. Secondly, the PEF scaling approach leads to a smaller number of scales that can be used than when the data are scaled. The PEF must be scaled by an integer value, so that the number of scales is constrained by the size of the data divided by the size of the PEF to be estimated. Conversely, if the data are scaled, this restriction is not present.

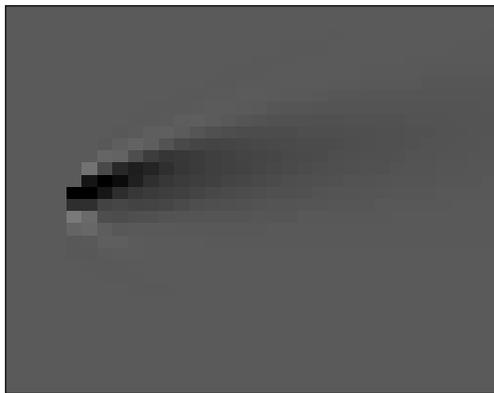
A simpler alternative would be to use a single scale of data, where there are sufficient fitting



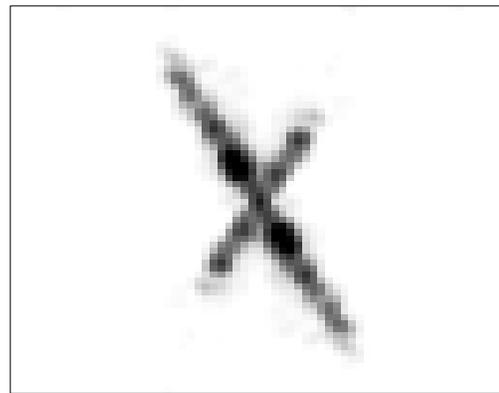
Original Data



Subsampled Data



PEF of Original Data



FT of Original Data

Figure 2: Clockwise from top left: Fully sampled test data; Sparsely sampled test data; Envelope of Fourier Transform of fully sampled test data, Inverse impulse response of PEF from all data. [bill1-data](#) [ER]

equations to adequately constrain the estimation of the PEF. Unfortunately, while the PEF is theoretically scale-invariant, there is some variation in the estimated PEF from scale to scale, which makes the choice of scale a challenge, as shown in Figure 3. Also, using multiple scales allows for more fitting equations, which will better constrain the estimation and minimize the effect of erratic data.

There are several degrees of freedom with the multiscale approach. First of all, the size of the PEF can be changed to accommodate a range of situations. A two-column PEF can annihilate a single plane wave, a three-column PEF can annihilate two plane waves, and so on. The height of the PEF determines the dip of plane waves that can be annihilated. Also, the size of the PEF also determines the number of fitting equations for each PEF, with a larger PEF meaning not only that each coefficient is less constrained, but that overall there are less fitting equations for the entire system, due to edge effects and more equations containing missing data. In this case a 5x3 PEF was chosen, so that it can eliminate two plane waves with dips ranging from $+63^\circ$ to -90° .

Another degree of freedom is the choice of scales. Certain large scales do not have enough equations to adequately constrain the PEF, and the PEF becomes unstable. Furthermore, at large scales the normalized linear interpolation returns distorted data, where the data are copied into nearby bins. Conversely, there becomes a point where the size of the PEF is approaching the size of the data, where the estimation of the PEF suffers from sampling issues and a lack of equations. Both of these cases are illustrated in Figure 3. In this example, the ranges were chosen to be from one half of the original scale to one quarter of the original scale, so that the bins would be adequately filled without the detrimental effects of normalized linear interpolation at the larger scales, and sampling issues at the lower scales. The scales which are used are ultimately a function of the sparseness of the data as well as the size of the PEF.

The result of the multiscale PEF estimation is not perfect, as shown in Figure 4. However, the estimate does contain the relevant dips, so it can reliably be used as a starting point for the nonlinear scheme, described next.

NONLINEAR PEF ESTIMATION

Once an initial estimate of the PEF is made, this estimate can be used to interpolate the data, from which a new PEF can be estimated. This nonlinear approach can be repeated until it converges to a final solution, shown in Figure 5. Unfortunately, like most nonlinear methods, the choice of a starting guess is crucial to the success and efficiency of the method. When the starting interpolation is far from the ideal solution, convergence to the best solution is not likely. Figure 6 shows the nonlinear method used on the data, with various different original guesses.

The nonlinear method appears to create wildly different solutions depending on the starting guess. Both the Laplacian interpolation as well as one of the single scale PEF interpolations produced poor results. Both the multiscale PEF and the other single scale PEF produced pleasing results, which are both very close to the ideal solution. The ideal solution is where the PEF

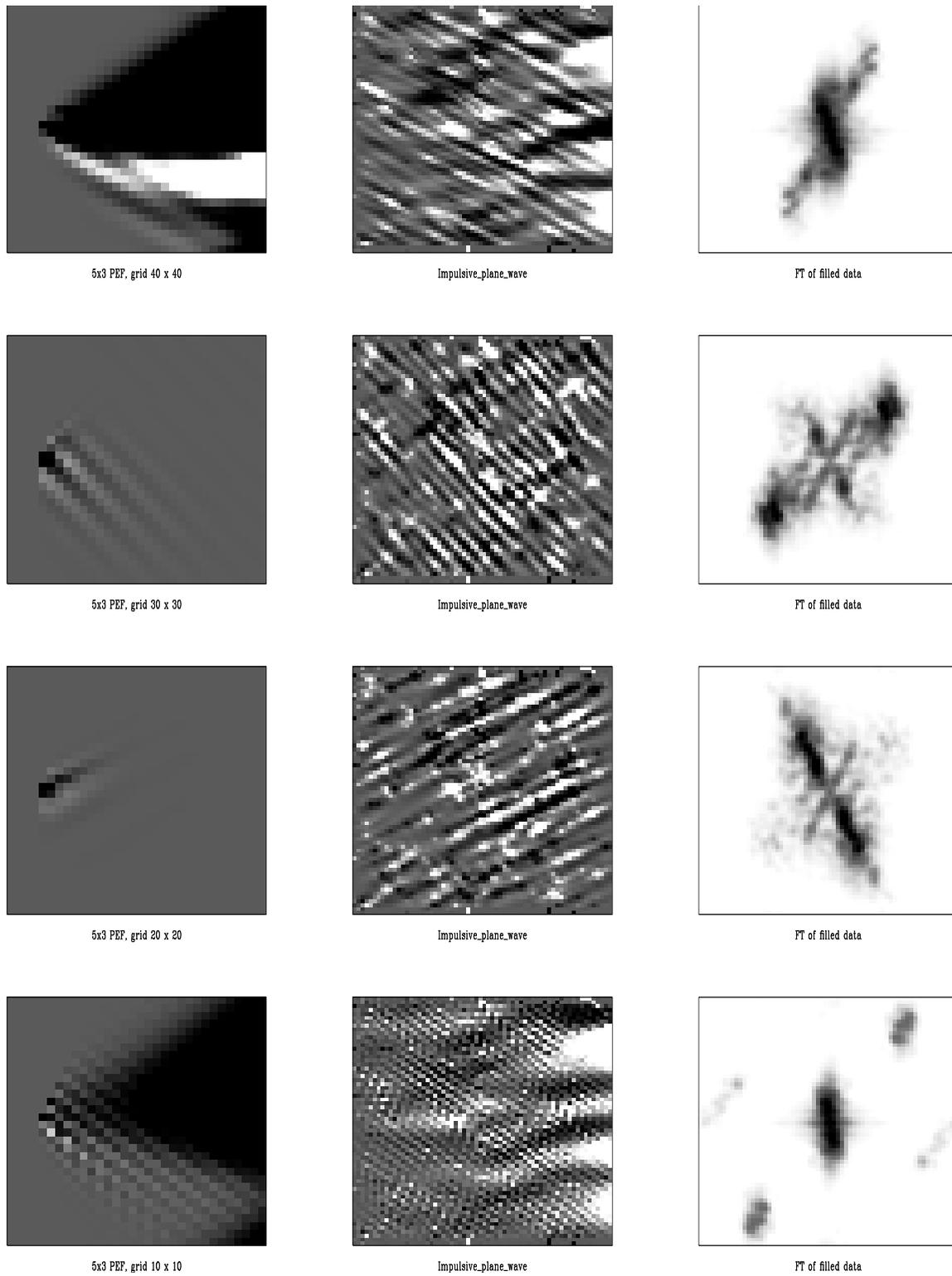


Figure 3: From left to right: Regridded data; inverse impulse response of estimated PEF; envelope of Fourier Transform of data filled with estimated PEF. From top to bottom: 40x40 grid; 30x30 grid; 20x20 grid; 10x10 grid bill1-refpefs [ER,M]

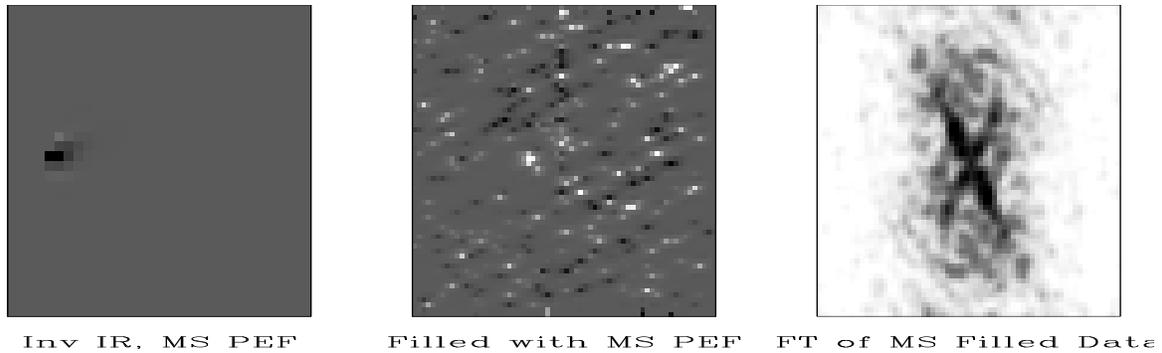


Figure 4: Multiscale estimation, from left to right: Inverse impulse response of PEF; data filled with PEF; envelope of Fourier Transform of filled data. The PEF was calculated with 9 scales of data ranging from $1/2$ of the original scale to $1/4$ of the original scale. `bill1-multipef` [ER]

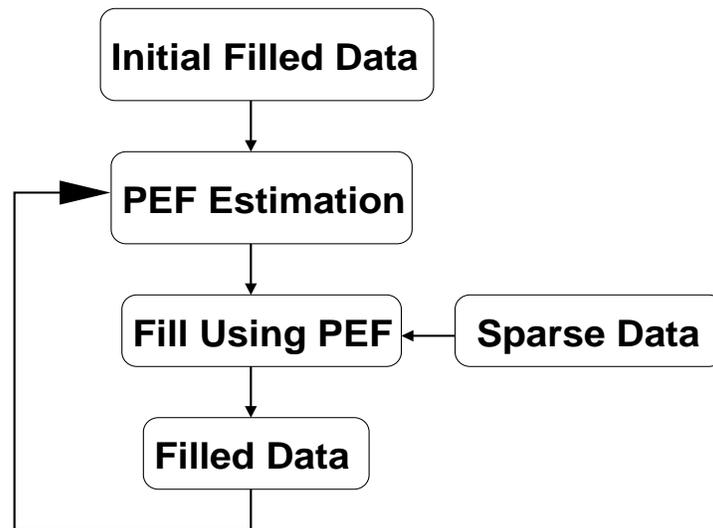


Figure 5: Flow chart for nonlinear PEF estimation. `bill1-nlflow` [NR]

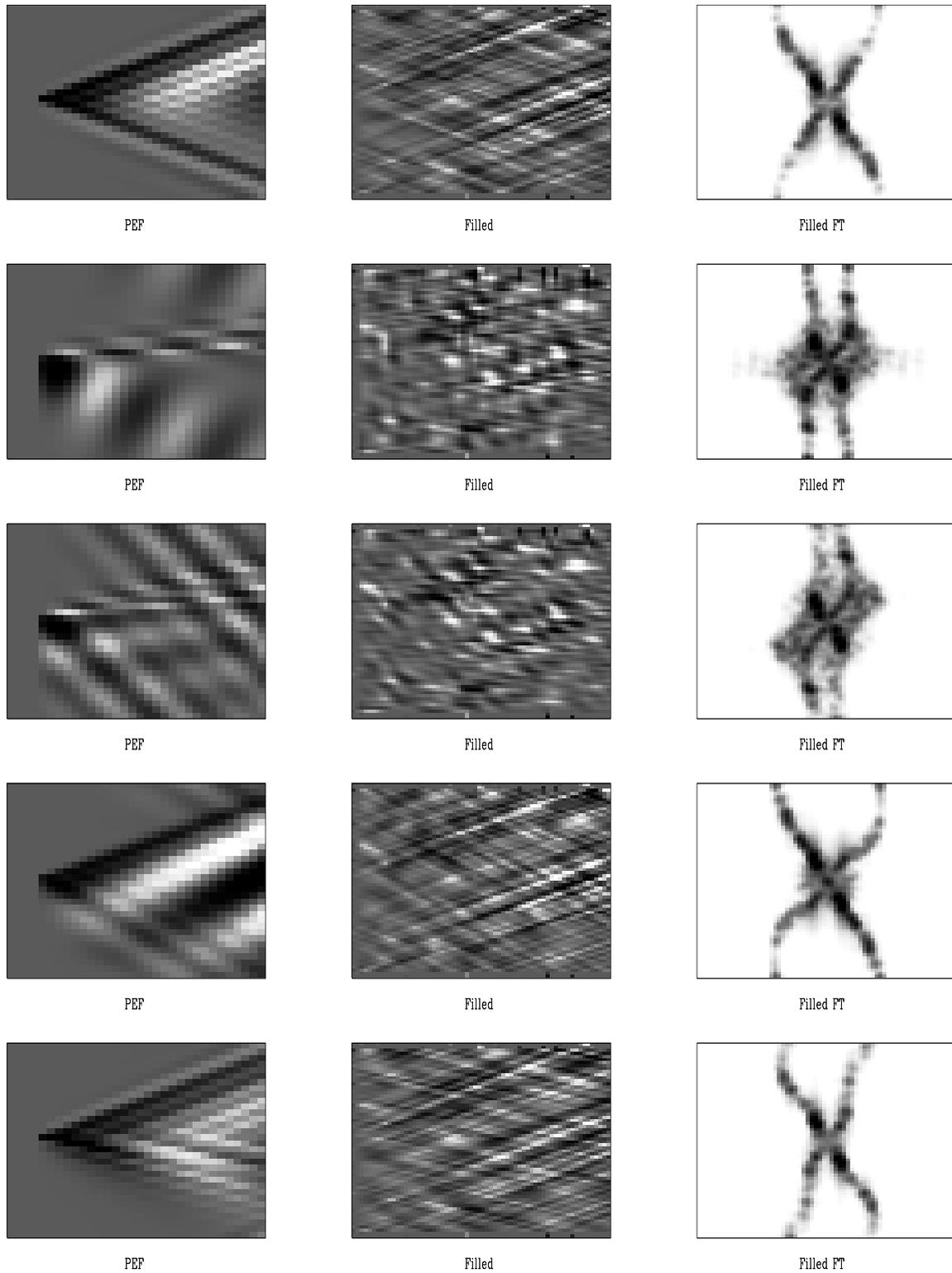


Figure 6: Nonlinear PEF estimation with various starting guesses. From left to right: inverse impulse response of PEF; filled data; envelope of Fourier Transform of filled data, all after 10 nonlinear iterations. From top to bottom: filled with PEF obtained from original data; filled with Laplacian; filled with single scale (32x32) PEF, filled with second single scale (20x20) PEF, filled with multiscale PEF. `bill1-nlfigure` [ER,M]

is calculated from the fully sampled data, and is then used as a starting guess in the nonlinear estimation. The multiscale result was obtained without knowing the solution, while the second single scale result was obtained by comparing various results with the original filled data, and selecting the most similar.

CONCLUSIONS

The estimation of a PEF with conventional methods for this example was impossible. Single scale PEF estimation can provide a good solution, but small changes in scale can create very different final solutions, especially when used as the starting guess for a nonlinear problem. Multiscale PEF estimation overcomes the variability in scales, and produces a result which is not perfect, but gives a robust result that requires less prior knowledge.

At this point, there are several avenues that can still be explored, such as examining the differences between normalized linear interpolation and binning while regridding, using weighting functions during both the multiscale PEF estimation and the nonlinear estimation, and developing better automated methods of choosing appropriate scales for the estimation.

REFERENCES

- Brown, M., Claerbout, J., and Fomel, S., 2000, Test case for PEF estimation with sparse data ii: SEP-105, 117–122.
- Claerbout, J., 1998, Multidimensional recursive filters via a helix: SEP-97, 319–335.
- Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, <http://sepwww.stanford.edu/sep/prof/>.
- Crawley, S., 1998, Shot interpolation for radon multiple suppression: SEP-97, 173–182.
- Crawley, S., 2000, Seismic trace interpolation with nonstationary prediction-error filters: Ph.D. thesis, Stanford University.
- Fomel, S., 2001, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.

Estimation of systematic errors in tracked datasets using least squares crossing point analysis

Morgan Brown¹

ABSTRACT

Geophysical data often contains both systematic and random errors. If left unchecked, the systematic errors can cause acquisition footprint in the final map. I present a method to estimate systematic error by analyzing measurements at points where two acquisition swaths cross. I then subtract the estimated systematic error from the data and generate a map with a familiar least squares formulation. I test the method on bathymetric data from the Sea of Galilee. Compared to two previous least squares formulations, my new method produces final maps which are relatively free of acquisition footprint, and which exhibit preservation of underlying bathymetric features.

INTRODUCTION

Simple tracked datasets, consisting of single-channel measurements (z) acquired by a moving instrument at surface points (x, y) , have traditionally proven valuable test-beds for least squares estimation techniques at SEP, due to their small size and conceptual simplicity. Ben-Avraham et al.'s (1990) bathymetric survey of the Sea of Galilee (Lake Kinneret) has drawn a prolonged interest (Claerbout, 1999; Fomel and Claerbout, 1995; Fomel, 2001), primarily because errors in the data seriously inhibit the task of translating the data into a gridded map.

In practice, each point measurement contains random errors, due both to instrumental inaccuracy and to physical phenomena with time and spatial scales smaller than two neighboring measurements. Unfortunately, these “random” errors are often non-gaussian, violating a central assumption of estimation theory that data contain gaussian-distributed error. In the context of Galilee, the ship's pre-GPS radio location system often mis-positioned depth soundings on the earth's surface. The erroneous measurements are easily identified as spikes in locations where the true sea floor is nearly flat, but not where the sea floor dips steeply. Even more crucially, however, these data also contain systematic error, which I define as error which varies slowly—in time and space—over a single track of measurements. Causes may include tidal shifts, instrumental drift, and wind-induced bulging of the lake's surface.

(Claerbout, 1999) cast the translation of irregular point data into a regular gridded map as an “inverse interpolation” problem. The simplest 2-D formulation of this problem is sensitive both to non-gaussian and systematic errors as noted by Fomel and Claerbout (1995). To

¹email: morgan@sep.stanford.edu

overcome both difficulties, they included a composite residual weighting operator consisting of: 1) diagonal weight estimated via Iteratively Reweighted Least Squares (IRLS) (Nichols, 1994; Guitton, 2000) to handle non-gaussian noise, and 2) a finite-difference first derivative operator to suppress correlated components of the residual. A similar technique was applied successfully to process a Geosat dataset with similar errors by Ecker and Berlioux (1995).

While Fomel and Claerbout's (1995) approach suppresses both acquisition footprint and non-gaussian noise, the authors note a loss of resolution in the final map of Galilee, relative to the simple inverse interpolation result. As noted by Claerbout (1999), when the systematic error varies from one acquisition track to another, a bank of prediction error filters, one for each acquisition track, makes a far better residual decorrelator. Karpushin and Brown (2001) implement this approach and report good suppression of acquisition footprint, with preservation of underlying bathymetric features.

In this paper, I take a somewhat different tack to the problem. I measure the difference in measured sea floor depth at "crossing points", or points in space where two acquisition tracks nearly cross. I then solve a least squares missing data problem to estimate the systematic error at all points between the crossing points, subtract the estimated systematic error from the original data, and use the simple inverse interpolation methodology of Claerbout (1999), supplemented with IRLS diagonal residual weights to suppress spikes, in order to make a final map. My approach generates maps of Galilee's seafloor which are generally free of acquisition footprint, and which also exhibit excellent preservation of subsea geologic features. Furthermore, my new approach visibly unbiases the residual, proving that the result is optimal from the standpoint of estimation theory.

METHODOLOGY

The simplest inverse interpolation approach outlined in (Claerbout, 1999) can be written in least squares fitting goals as follows.

$$\begin{aligned} \mathbf{B}\mathbf{m} - \mathbf{d} &\approx \mathbf{0} \\ \epsilon \mathbf{A}\mathbf{m} &\approx \mathbf{0} \end{aligned} \quad (1)$$

\mathbf{B} is nearest neighbor interpolation and maps a gridded model (\mathbf{m}) to the irregular data space (\mathbf{d}). The model grid is 860x500 points, while the data space consists of over 132,000 (x, y, z) triples. \mathbf{A} is a model regularization operator, which penalizes model roughness. For all examples contained herein, $\mathbf{A} = \nabla$. ϵ balances the tradeoff between data fitting and spatial model smoothness.

To handle non-gaussian noise, Fomel and Claerbout (1995) implement an Iteratively Reweighted Least Squares (IRLS) scheme to nonlinearly estimate a residual weight which automatically reduces the importance of "bad" data in least squares estimation. Adding a diagonal residual weight to equation (1) gives

$$\begin{aligned} \mathbf{W}(\mathbf{B}\mathbf{m} - \mathbf{d}) &\approx \mathbf{0} \\ \epsilon \mathbf{A}\mathbf{m} &\approx \mathbf{0}. \end{aligned} \quad (2)$$

To handle systematic errors between data tracks, Fomel and Claerbout (1995) supplement the residual weight in system (1) with a first derivative filter to decorrelate the residual. Lomask (1998) used a single prediction error filter (PEF). Karpushin and Brown (2001) use a bank of PEF's, one for each acquisition track. Whatever the case, we can refer to the differential operator as \mathbf{D} and modify equation (2) to obtain a new system of equations:

$$\begin{aligned}\mathbf{WD}(\mathbf{Bm} - \mathbf{d}) &\approx \mathbf{0} \\ \epsilon\mathbf{Am} &\approx \mathbf{0}.\end{aligned}\tag{3}$$

\mathbf{W} is the same as in equation (2), except for the addition of zero weights at track boundaries.

Estimation of systematic error from crossing points

It is fairly easy to compute “crossing points” from the raw data. I partition the data spatially into fairly small regions and, using a Fortran90 data structure, store the track index of each data point in the region. I decide that two data points cross if they belong to different tracks and are separated (spatially) by a distance less than a predefined threshold.

Unraveling the distribution of errors between the two crossing points is a much more difficult, and possibly intractable problem, in the absence of prior information. In general, each of the two crossing points contain an unknown combination of systematic and random error. In this case, I simply “split the difference” by assuming that, given a measured difference between two crossing points, the systematic error is distributed evenly between the two points.

Once I have a series of measured differences at crossing points, I estimate the systematic error at all points in the data space by solving a least squares missing data problem. I impose the assumption that the systematic error is smooth along acquisition tracks by adding a roughness penalty to the least squares objective function. In least squares fitting goals, these ideas are written as follows.

$$\begin{aligned}\mathbf{W}_e(\mathbf{Ke} - \mathbf{Ke}_d) &\approx \mathbf{0} \\ \nabla\mathbf{e} &\approx \mathbf{0}\end{aligned}\tag{4}$$

\mathbf{e}_d is the measured differences at crossing points. \mathbf{e} is the estimated systematic error. The reliability of the measured difference is inversely proportional to the spatial separation distance of the two crossing points, r . Arbitrarily, I choose the residual weight, \mathbf{W}_e , for a pair of crossing points as $(1 - (r/r_0)^2)$, where r_0 is the predefined threshold which defines crossing points. \mathbf{K} is a selector matrix. Figure 1 outlines the scheme.

Why Bother Directly Estimating Systematic Error?

If system (3) produces track-free maps, the skeptic might wonder, *why even bother directly estimating the systematic error in the data?* First, as mentioned earlier, system (3) leads to a loss of resolution in the final map. Second, in many cases we may have prior information

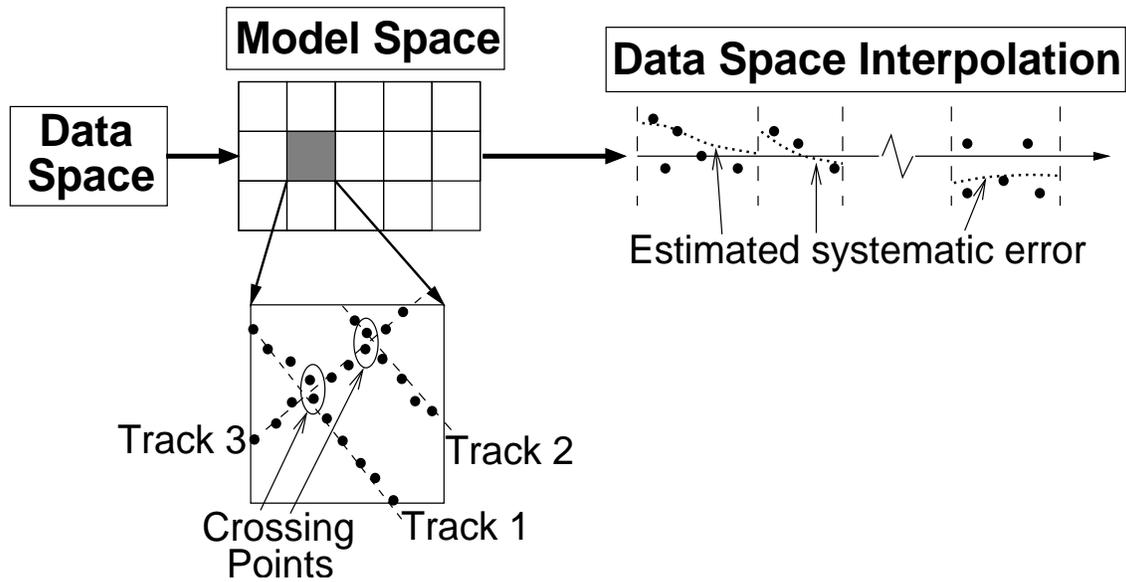


Figure 1: Algorithm for estimating systematic data error by inverting crossing point data. Data is partitioned into local regions, where crossing point differences are computed. The optimization of equation (4) operates in the data space. The measured crossing point differences are the data; we use least squares to fit a smooth trend to the measured data. The optimization operates independently on individual tracks. `morgan1-crosspoint` [NR]

about the distribution and magnitude of the systematic error, which we could then include as an “inverse model covariance” (regularization operator) in system (4).

By directly estimating and subtracting systematic errors, we have more faith that the final map is an accurate representation of the true quantity. While the authors of the previously mentioned papers on Galilee and Madagascar were more interested in resolving the topographical features than the value of the underlying field, in many applications, the field itself is most important. Furthermore, when the systematic error is explainable by physical or other phenomena, we want to have control in its estimation.

Building a map with my new approach

My new approach can be summarized simply:

1. Estimate systematic error in data via system (4).
2. Subtract estimated systematic error from original data.
3. Generate final map using system (2).

RESULTS

In this section, I compare the results on the Galilee data of using systems (2) and (3) with my approach. I find the raw final maps uninterpretable, so instead, I will show roughened (first derivative) versions of the final maps, as well as the differences between the three final maps.

Figure 2 shows the final Galilee maps computed with the three discussed methods, and then roughened from southwest to northeast. The block-shaped artifacts on the periphery of the study area are due to the “Quadtree Pyramid” method used to generate a starting guess for the conjugate gradient iteration (Brown, 2000).

The result for system (3) in Figure 2 is free of the acquisition footprint which plagues the center panel (system (2)), but obscured by seemingly random noise. This noise is a result of track-end artifacts which “propagate”, even though we have applied a zero weight on both sides of track boundaries. We could suppress the random noise by increasing ϵ , at the expense of resolution. The right panel of Figure 2 shows the result using my approach. Track artifacts are suppressed considerably, though not totally, compared to the center panel. The underlying geologic features look just as well resolved as in the center panel.

Figure 3 shows the difference between final maps generated by each of the three methods discussed in this paper. The left panel shows the difference between system (3) and system (2). We see track artifacts, as well as geologic features (238km north, 205km east), which confirms that first derivative along the track added to system (3) has caused some loss of resolution. The center panel of Figure 3 is the difference between the system (3) map and the map generated by my approach. First, we notice again the same loss of geologic features as in the left panel. Interestingly, we see correlated differences that are not along tracks. This tells us that although both approaches lead to white residuals, as we will see in Figure 4, they do not produce identical maps. This question needs to be answered. The right panel of Figure 3 shows the difference between my approach and system (2). As expected, we see considerable differences along tracks, but little to no geologic differences. This tells us that my approach has maintained the resolution inherent in system (2), while doing a good job of suppressing acquisition footprint.

In the northern region of the map, we notice mainly negative differences along east-to-west tracks, and positive differences along north-to-south tracks. In the southern region, this relationship is reversed, and the differences are noticeably smaller. The spatial regularity in differences implies that the systematic errors may well be correlated in time and/or space. Unfortunately, we do not have the times at which the samples were collected. We only know that they were collected between 1986 and 1987.

Figure 4 compares the data residuals for each of the three discussed methods. First, we see that IRLS alone (system (2)) produces a biased residual. The bias is most easily seen around sample 60,000 and sample 110,000. IRLS with track derivative (system (3)) produces an unbiased residual. On the bottom panel of Figure 4, we see that my approach leads to an effectively unbiased residual. Thus, from the viewpoint of optimization theory, my approach has achieved one half of the most important requirement of an optimal map: it is unbiased, although by inspection, not white. A better IRLS scheme should produce a more balanced,

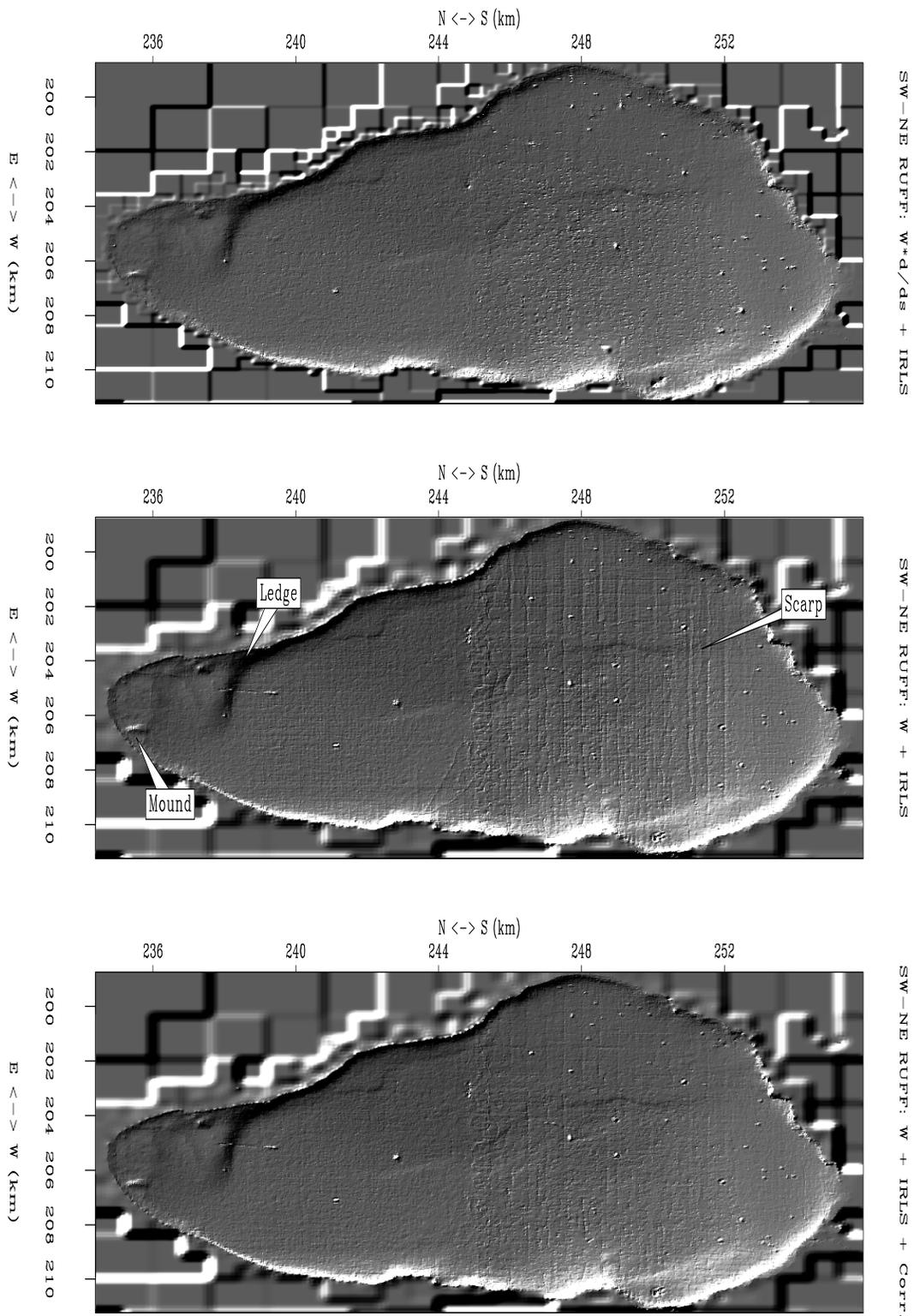


Figure 2: Final Galilee maps, roughened from southwest to northeast, using a simple difference filter. Left: IRLS with track derivative (system (3)). Center: IRLS only (system 2)). Right: My new approach—IRLS + crossing point correction. [morgan1-galilee-ruff2](#) [ER,M]

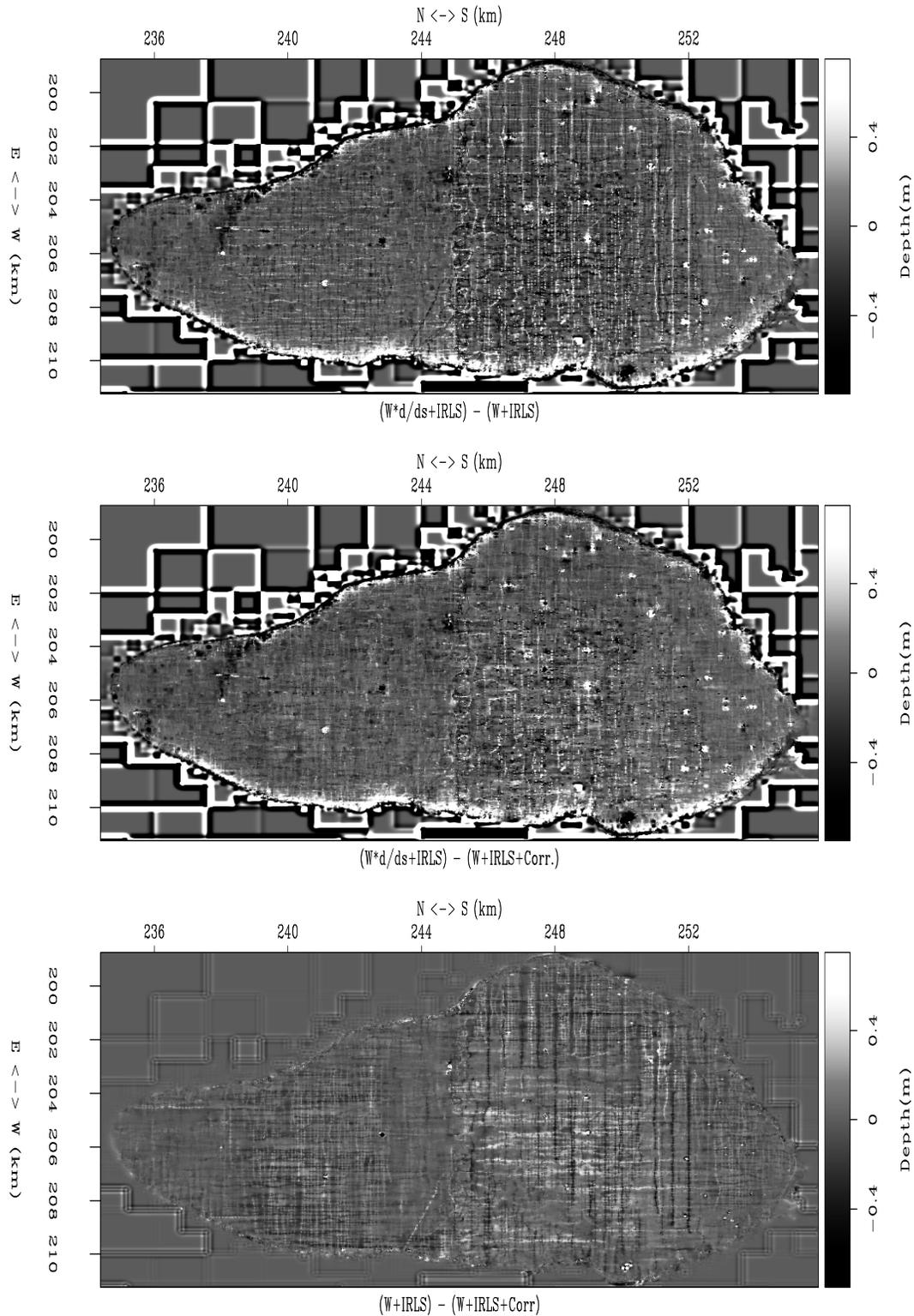


Figure 3: Difference between final Galileo maps produces by three discussed methods. Top: IRLS with track derivative (system (3)) minus IRLS only (system (2)). Center: IRLS with track derivative minus my new approach—IRLS + crossing point correction. Bottom: IRLS only minus my new approach. `morgan1-galilee-diff` [ER]

and hence white, residual for all three approaches.

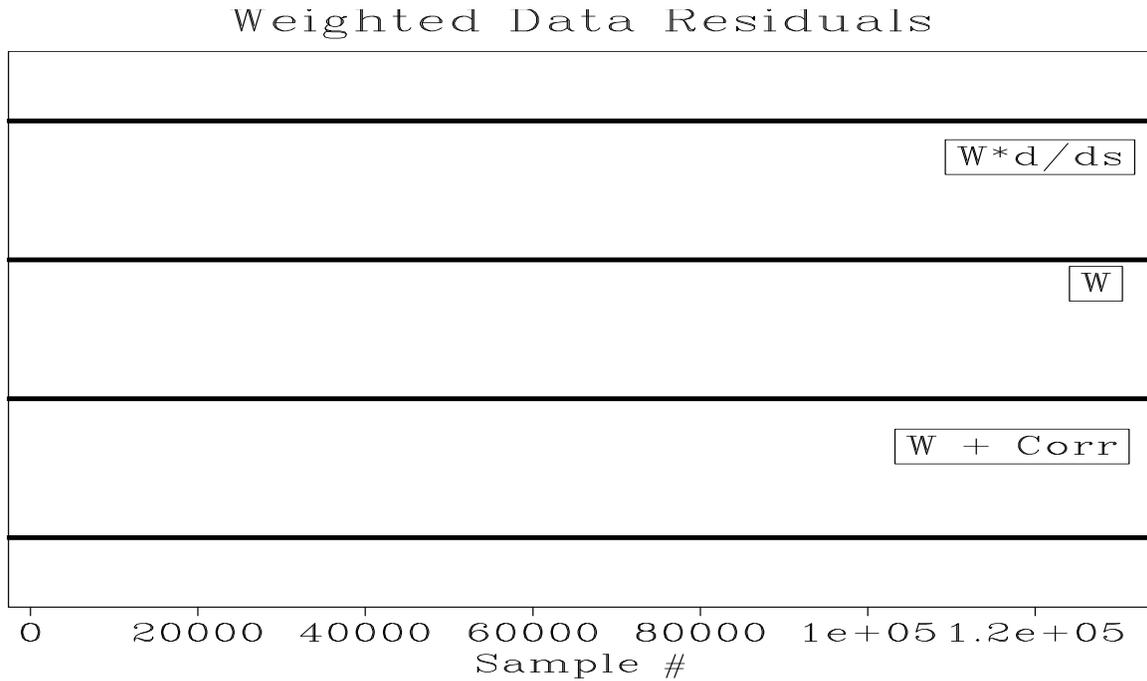


Figure 4: Residual error for the three discussed methods. Top: IRLS with track derivative (system (3)). Center: IRLS only (system (2)). Bottom: My approach. `morgan1-galilee-resid [ER]`

DISCUSSION

I presented a new method for building maps from tracked datasets. As discussed in previous works dealing with the Galilee bathymetry data, the biggest problem in building a map is systematic error between neighboring acquisition tracks. My method estimates systematic error between tracks by directly analyzing the difference between tracks at “crossing points” and using least squares optimization to estimate the error in points in between.

As we saw in tests on the Galilee data, my method effectively unbias the data residual without a loss of resolution. My approach produces the most interpretable Galilee maps of the three approaches I discussed. Although the track artifacts are not removed completely, I believe the increased resolution relative to IRLS + track derivative approach weighs the balance in my approach’s favor.

Looking to the future, I believe my approach offers a fundamental advantage over those of system (3)’s ilk. The philosophy behind system (3) assumes that we know that the residual is biased, but that we don’t necessarily understand the data errors that cause the bias. In many cases, we *do* have strong prior information on the cause, magnitude, physics, and covariance of the systematic error. A good estimate of the systematic error may have interpretive value.

One weakness of my approach is that it reeks of “preprocessing”. A more general and rigorous approach is that of Brown and Clapp (2001), which is basically an iterative variant of the approach that I’ve proposed here. Unfortunately, without prior information on the character of the systematic error, such an approach would be wasted.

REFERENCES

- Ben-Avraham, Z., Amit, G., Golan, A., and Begin, Z., 1990, The bathymetry of Lake Kinneret: Israel J. Earth Sci, **39**, 77–84.
- Brown, M., and Clapp, R. G., 2001, A least-squares approach for estimating integrated velocity models from multiple data types: SEP-**108**, 159–172.
- Brown, M., 2000, Two strategies for sparse data interpolation: SEP-**105**, 49–66.
- Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, <http://sepwww.stanford.edu/sep/prof/>.
- Ecker, C., and Berlioux, A., 1995, Flying over the ocean southeast of Madagascar: SEP-**84**, 295–306.
- Fomel, S., and Claerbout, J., 1995, Searching the Sea of Galilee: The splendors and miseries of iteratively reweighted least squares: SEP-**84**, 259–270.
- Fomel, S., 2001, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.
- Guitton, A., 2000, Huber solver versus IRLS algorithm for quasi L1 inversion: SEP-**103**, 255–271.
- Karpushin, A., and Brown, M., 2001, Whitening track residuals with PEFs in IRLS approach to the Sea of Galilee: SEP-**110**, 133–140.
- Lomask, J., 1998, Madagascar revisited: A missing data problem: SEP-**97**, 207–216.
- Nichols, D., 1994, Velocity-stack inversion using L_p norms: SEP-**82**, 1–16.

Whitening track residuals with PEFs in IRLS approach to the Sea of Galilee

Andrey Karpushin and Morgan Brown¹

ABSTRACT

We applied an Iteratively Reweighted Least Squares (IRLS) approach to create a map of the Sea of Galilee. We use a bank of Prediction Error Filters (PEFs) as a residual whitener to reduce the acquisition footprint and map artifacts caused by non-Gaussian noise in the data.

INTRODUCTION

The bottom sounding survey data set collected by Zvi Ben-Avraham in 1986-87 on the Sea of Galilee in Israel (Ben-Avraham et al., 1990) is often used at SEP to test new algorithms for data interpolation. The main difficulties associated with interpolating the Galilee data set onto a regular grid are inconsistencies in the data due to random spikes, and an acquisition footprint (ship tracks) left in the final image caused by systematic error in the data.

There are different approaches developed at SEP to confront these difficulties (Claerbout, 1999; Fomel and Claerbout, 1995; Fomel, 2001; Brown, 2001). Fomel and Claerbout (1995) suggested to use a derivative operator to filter out low-frequency components of the residual, and the Iteratively Reweighted Least Squares (IRLS) technique to suppress non-Gaussian spikes in the data. However, they showed that while the acquisition footprint and the noisy portion of the model disappear, the price of the improvement is a loss of image resolution.

In this paper we implement an idea from Claerbout (1999), and use a bank of a Prediction Error Filters instead of the derivative operator to whiten the residual on individual data tracks before implementing IRLS. We show that our algorithm leads to reduced artifacts in the final map without an apparent decrease in map resolution.

DATA DESCRIPTION

The Galilee data set includes more than 131,500 (x_i, y_i, z_i) triples where x_i varies over about 12 km and y_i varies over about 20 km. The water depth was measured using an Odom Ech-track DF3200 Echosounder, and the position of the boat at the time of the measurement was determined using a Motorola Miniranger system, which is composed of radio stations on shore

¹email: andrey@sep.stanford.edu, morgan@sep.stanford.edu

and a recorder on the boat. Navigation was recorded using “event marks” rather than time, and no information on time is available. Although the precision of the acquisition system is high, there are some defects in the data caused by an occasional malfunction of the equipment. This creates artifacts in the final map, and makes processing of this dataset a challenging problem. Another source of artifacts in the model is the changing conditions of the acquisition, which create a strong footprint in the final map.

METHOD

The problem of interpolating irregularly sampled data, like the Galilee data set, onto a regular grid to produce a map can be written in terms of the fitting goals of an inverse linear interpolation problem (Claerbout, 1999):

$$\begin{aligned}\mathbf{0} &\approx \mathbf{Lm} - \mathbf{d} \\ \mathbf{0} &\approx \epsilon \mathbf{Am}\end{aligned}\tag{1}$$

\mathbf{L} is the linear operator which maps data onto the map. Usually \mathbf{L} is either a binning or a bi-linear interpolation operator. The second equation in system (1) is a regularization term where \mathbf{A} is a roughening operator that imposes smoothness of the model in this underdetermined problem, at the price of fitting data exactly. Everywhere below, \mathbf{L} is a binning operator (\mathbf{B}) and \mathbf{A} denotes a gradient filter consisting of two first order derivatives.

To suppress the artifacts caused by non-Gaussian noise in the data, Fomel and Claerbout (1995) introduced a weighting operator \mathbf{W} :

$$\begin{aligned}\mathbf{0} &\approx \mathbf{W}(\mathbf{Bm} - \mathbf{d}) \\ \mathbf{0} &\approx \epsilon \mathbf{Am}\end{aligned}\tag{2}$$

The choice of the weighting operator \mathbf{W} follows two formal principles:

1. Statistically bad data points (spikes) are indicated by large values of the residual $\mathbf{r} = \mathbf{Bm} - \mathbf{d}$
2. Abnormally large residuals attract most of the conjugate gradient solvers effort, directing it the wrong way. The residual should be whitened to distribute the solvers attention equally among all the data points to emphasize the role of the “consistent majority”

Based on these principles operator \mathbf{W} in equation (2) was chosen by Fomel and Claerbout (1995) to include two components: the first derivative filter \mathbf{D} taken in the space along the record tracks and the diagonal weighting operator $\tilde{\mathbf{W}}$.

$$\begin{aligned}\mathbf{0} &\approx \tilde{\mathbf{W}}\mathbf{D}(\mathbf{Bm} - \mathbf{d}) \\ \mathbf{0} &\approx \epsilon \mathbf{Am}\end{aligned}\tag{3}$$

Fomel and Claerbout (1995) chose weighting function to be $\tilde{\mathbf{W}}_i = \tilde{\mathbf{W}}(\mathbf{r}_i) = \frac{2\bar{r}}{|\mathbf{r}_i| + \bar{r}_i}$, where \bar{r} stands for the median of the absolute values from the whole dataset, and \bar{r}_i is the median in

a small window around a current point r_i . Since $\tilde{\mathbf{W}}$ depends upon the residual, the inversion problem becomes non-linear and system (3) can be solved using a piece-wise linear approach (Fomel and Claerbout, 1995). However, Fomel and Claerbout (1995) showed that while the noisy portion of the model disappeared, the price of the improvement is a loss of the image resolution.

In our approach we use a bank of PEFs to decorrelate the residual. Using Prediction Error Filters as a residual whitener better satisfy the second of the formal principles used by Fomel and Claerbout (1995). Since the character of the systematic errors in the data may vary in time and upon the location of the ship, an individual PEF \mathbf{P}_i is estimated for each data track from the residual obtained after solving system (1). In this case we defined a data track as a series of measurements recorded with a distance less than 100 meters between consecutive data points. System (3) then becomes:

$$\begin{aligned}\mathbf{0} &\approx \tilde{\mathbf{W}}\mathbf{P}(\mathbf{B}\mathbf{m} - \mathbf{d}) \\ \mathbf{0} &\approx \epsilon\mathbf{A}\mathbf{m}\end{aligned}\quad (4)$$

\mathbf{P} is an operator composed from PEFs \mathbf{P}_i . After a bank of PEFs is estimated we solve this non-linear problem [equation (4)] in the manner of piece-wise linearization similar to Fomel and Claerbout (1995). The first step of the piece-wise linearization is the conventional least squares linearization. The next step consists of reweighted least squares iterations made in several cycles with reweighting applied only at the beginning of each cycle. We chose the weighting function $\tilde{\mathbf{W}}$ to be $\tilde{\mathbf{W}}_i = \tilde{\mathbf{W}}(\mathbf{r}_i) = \frac{1}{(1+r_i^2/r_i^2)^{1/4}}$ (Claerbout, 1999).

In the next section we compare the maps obtained by the three different methods.

RESULTS

Figure 1 shows a roughened image of the lake bottom constructed by using the three methods discussed above. All of the three models were computed with the same value of damping parameter (ϵ) in a linear step of the IRLS iterations to make a fair comparison.

The top panel of Figure 1 shows the result corresponding to system (2), implementing IRLS without whitening the residual. This panel has a strong acquisition footprint; some “spikes” are still present in the model, even though IRLS managed to attenuate most of them. At the same time this panel has a good resolution and we can see some fine features of the bottom, especially in the southern part of the lake where the acquisition footprint is weaker. The central panel of Figure 1 shows the result of solving system (3) with a derivative operator along the tracks as a residual whitener. This panel has no acquisition footprint but has a lower resolution as was observed by Fomel and Claerbout (1995). Although this method preserved main features of the bottom, some features such as a ledge marked in the image are not present in the model. The third panel presents the result of our algorithm. As we can see almost all of the ship’s tracks and spikes are suppressed without a visible decrease in resolution. For example, the ledge we observed before is still present in the model. It is interesting to notice that the scarp is even easier to see here than in the first image, where it is obscured by the acquisition footprint.

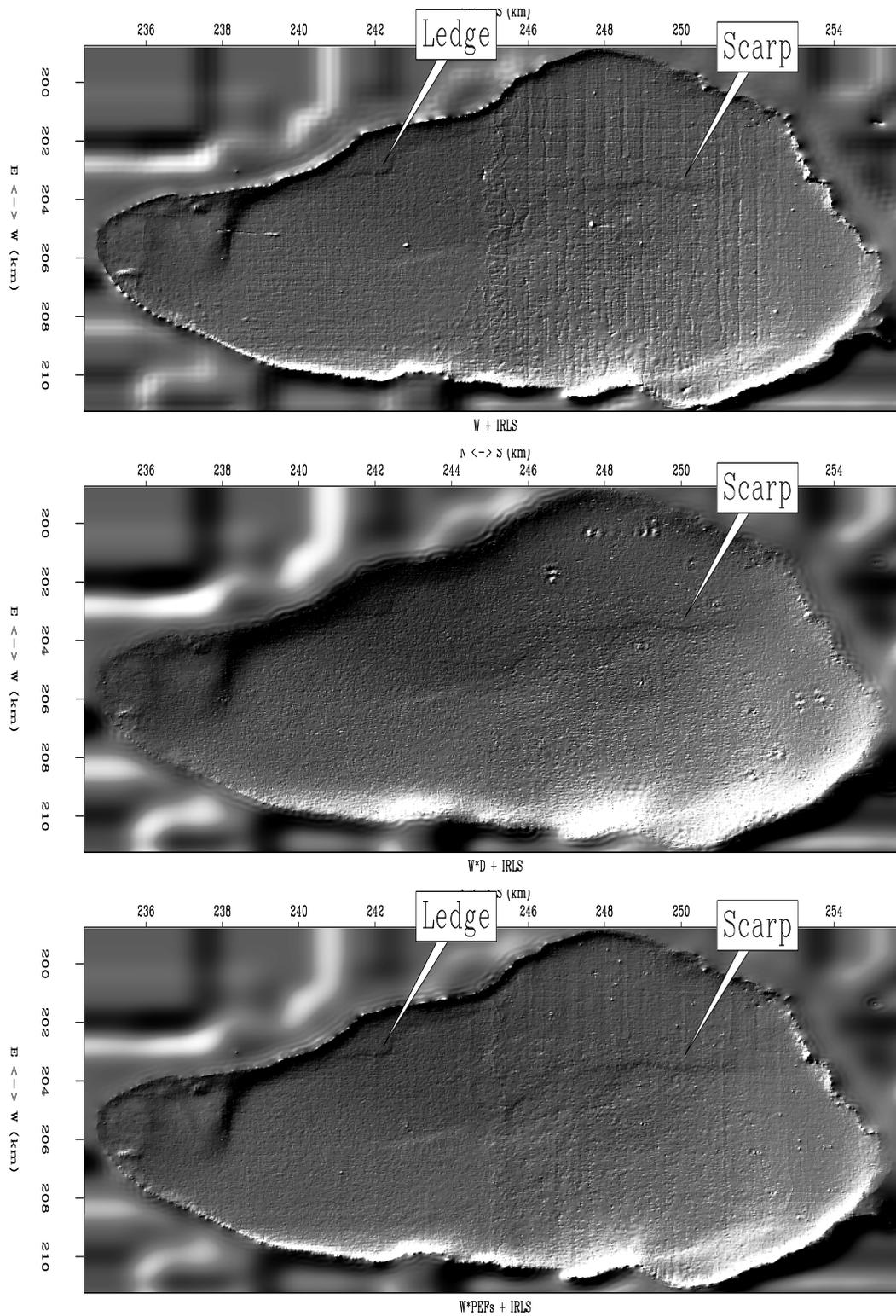


Figure 1: Comparison of the roughened results obtained by the three methods. Top: Result of IRLS with no residual whitener (system 2). Central: IRLS with a derivative operator as a residual whitener (system 3). Bottom: Our algorithm with a bank of PEFs as a residual whitener (system 4) `andrey-comp_1_ann` [ER]

Figure 2 shows the difference among the results computed with the three algorithms discussed above.

The top panel shows the difference between the maps computed with IRLS without whitening the residual and IRLS with a derivative operator as a residual whitener. On this panel we can see ship tracks and features of the bottom which were attenuated from the final map. The central panel shows the difference between the maps computed with IRLS without whitening the residual and IRLS with a bank of PEFs as a residual whitener. Geological features are much less visible in this panel, but it is easy to see the acquisition footprint that our method removed. The third panel shows the difference between the final maps generated with the use of a derivative and a bank of PEFs to whiten the residual. On this panel we can still see some traces of the acquisition footprint which were left in the final map by our method. It is also easy to see the geological features of the bottom of the lake which were preserved by our method.

CONCLUSIONS

We presented and tested a new method of interpolation of tracked datasets with non-Gaussian errors onto a regular grid. We showed that with the Galilee dataset, our method was able to remove artifacts caused by irregular spikes in the data as well as the acquisition footprint caused by systematic errors in the data. We think that this method is directly applicable to other problems where measurements are taken along the tracks and data have errors due to equipment malfunctioning and systematic errors due to slowly changing acquisition conditions.

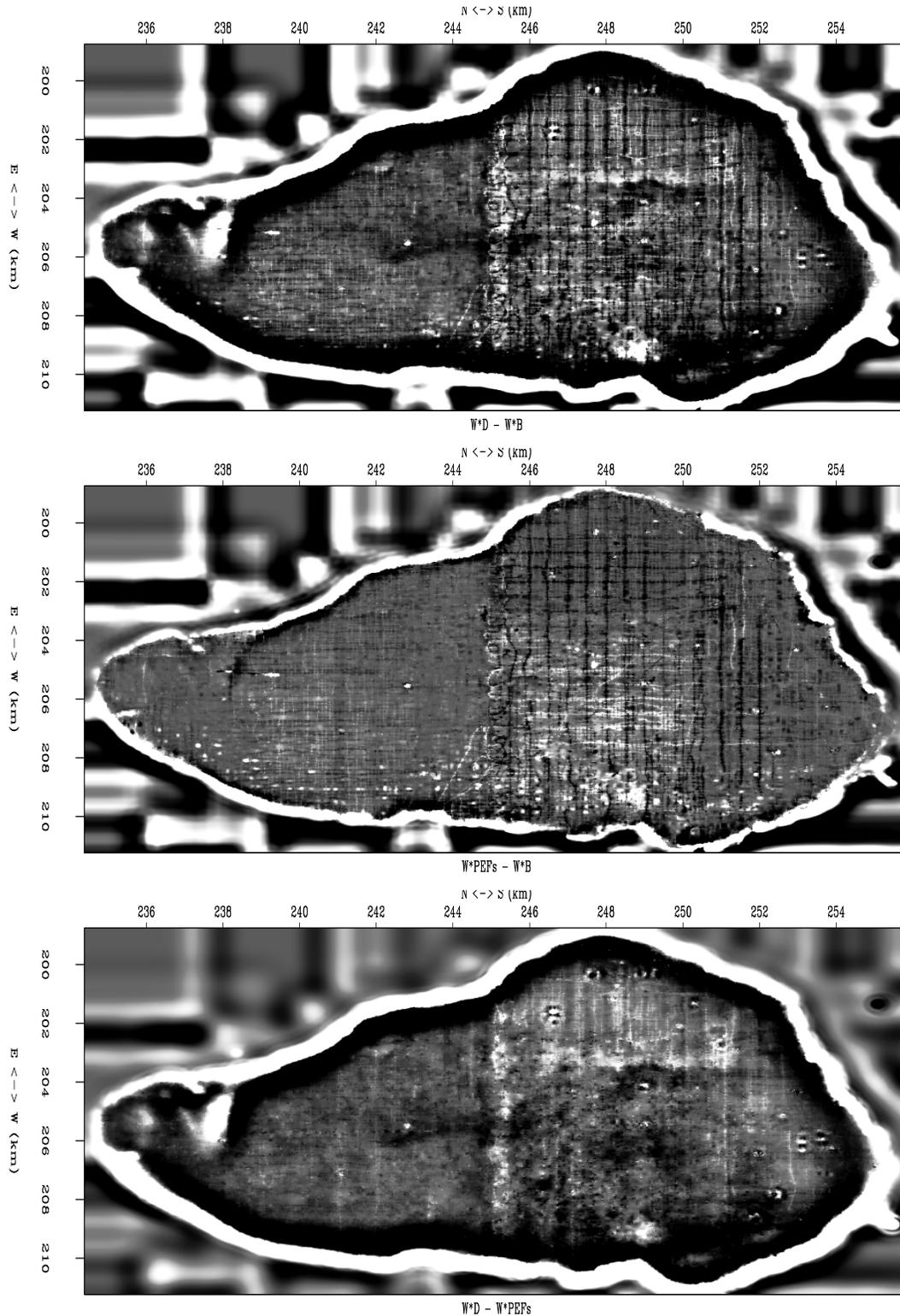


Figure 2: Difference in the results obtained by the three methods. Top: Difference between IRLS without whitening the residual and with a derivative operator as a whitener. Central: Difference between IRLS without whitening the residual and IRLS with a bank of PEFs as a residual whitener. Bottom: Difference between IRLS with a derivative and a bank of PEFs as a whitener. [andrey-comp_2](#) [ER]

REFERENCES

- Ben-Avraham, Z., Amit, G., Golan, A., and Begin, Z., 1990, The bathymetry of Lake Kinneret: *Israel J. Earth Sci.*, **39**, 77–84.
- Brown, M., 2001, Estimation of systematic errors in tracked datasets using least squares crossing point analysis: *SEP-110*, 123–132.
- Claerbout, J., 1999, Geophysical estimation by example: Environmental soundings image enhancement: Stanford Exploration Project, <http://sepwww.stanford.edu/sep/prof/>.
- Fomel, S., and Claerbout, J., 1995, Searching the Sea of Galilee: The splendors and miseries of iteratively reweighted least squares: *SEP-84*, 259–270.
- Fomel, S., 2001, Three-dimensional seismic data regularization: Ph.D. thesis, Stanford University.

Short Note

5th dimension warning for 4D studies

Brad Artman¹

MOTIVATION

It has been well documented that reservoirs show sometimes staggering amounts of compaction over years of production. The aquifer under Las Vegas, the Ecofisk development, and the Lake Maracaibo area are all examples where the subsidence of porous reservoirs have undergone a compaction observable to the naked eye.

Studies of the elastic deformation of the surrounding country rock have shown the distribution of failure types to be associated with the volumetric collapse of a reservoir structure (Segall, 1998). However, the prediction of the nature and timing of the compaction of the reservoir itself need also be understood. Zoback (2001) has made inquiries into the nature of the compaction of granites, Ottawa sand, and the Adamswiller sandstone. This work focuses on an elastic deformation model that does not take into account time progression and creep behavior.

The motivation for the study of the creep behavior of reservoir rocks is a fundamental problem of available time. This uncertainty manifests itself in two questions: Over what time scale do reservoirs compact? Can lab measurements capture meaningful parameters to understand the nature and development of reservoir compaction? Simply put, are compaction studies measured in the lab valid at all? After these difficult questions are answered, then next logical step is to ask what impact this can have on repeat seismic experiments that are all the rage in the new millennium.

With the publishing of lab compaction data in Dudley and Myers (1994), the case is clear that meaningful measurements can be made despite implementation on time scales orders of magnitude shorter than significant to either geologic time or reservoir life. This results in our ability to make and trust compaction measurements on unconsolidated sand material. While this may solve engineering problems such as pore volume compressability and drive, it introduces a major wrinkle into the 4D seismic experiment that must be recognized. As a reservoir continues to compact over a decade of production, the best models for fluid substitution or pressure dependence of a seismic attribute are meaningless if the rock frame has changed significantly.

¹email: brad@sep.stanford.edu

COMPACTION

The uniaxial compaction coefficient is defined as

$$C_z = \frac{d\epsilon}{d\sigma_z}. \quad (1)$$

Thus if the total amount of strain, ϵ , is different for tests in which the residence time of the stress, σ , varies we will measure a suite of compaction coefficients depending on our test parameters rather than inherent properties of the rock. Intuitively, we would expect this to happen; surely holding pressure constant for a longer period of time will result in more compaction.

To investigate this problem, several uniaxial strain experiments were measured by Dudley and Myers (1994) by increasing axial stress and holding for 1.5 hours, 1 day, 1 week and 1 year. The three shorter duration tests were repeated for 12 stress and hold cycles, while the year-long experiment was only conducted once. Figure 1 shows three such tests with 750 psi stress step. All samples are brine saturated and allow fluid to flow out of the sample with applied stress.

Figure 1: Apparently unrelated stress-strain behavior of samples to uniaxial strain experiment. (Dudley and Myers, 1994) `brad1-creep` [NR]

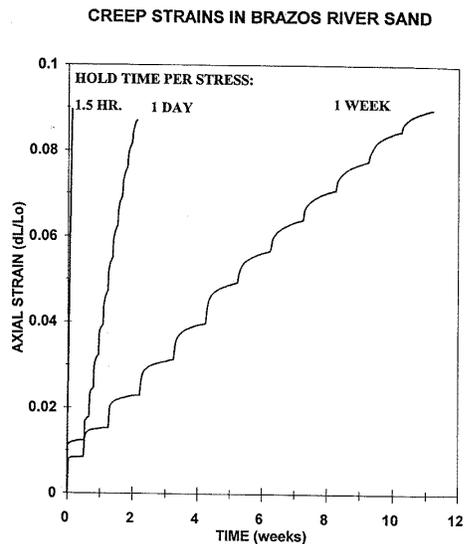


Figure 2 shows an example of the raw data from one stress cycle. Notice the extraordinary fit of the exponential curve to the integrated strain measurements over time. Figure 3 shows the surprising fact that by simply normalizing the time axis by the length of the hold time, the creep graphs all become coincident. This means that creep rate for longer experiments is slower than that for shorter experiments. This is simply amazing.

Several unsettling problems arise from these observations:

- Why does a growing exponential, that has no upper bound, fit data so well?

Figure 2: 24 hour hold time test for cycle to 8500 psi. Exponential fit of the form $\epsilon(t) = kt^d$ overlay the integral of the strain data. (Dudley and Myers, 1994) brad1-raw [NR]

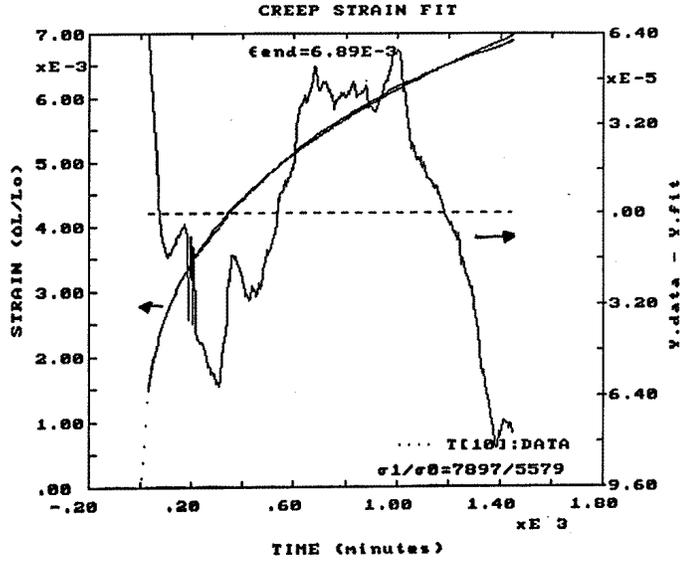
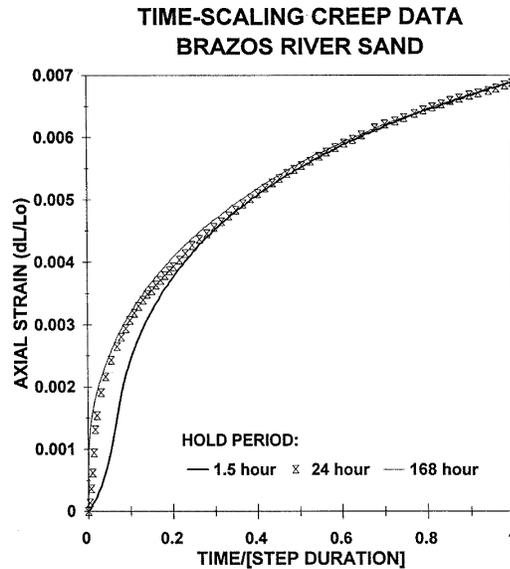


Figure 3: Amazingly, all three tests show the same exponential increase of strain as a function of normalized time. More important: all tests attain the same final strain value despite two orders of magnitude difference in residence time. Mismatch of 1.5 hour test assumed to be a result of time for 750 psi pressure ramp-up being a significant (13%) portion of total hold time. (Dudley and Myers, 1994) brad1-norm [NR]



- What does final strains being equal afford us?
- Why do normalized data overlay?

SEARCH FOR MODEL

Intuitively, we know that the creep behavior of the samples must reach a maximum that we can call *compressed*. The lack of an upper bound is troubling. Fortunately similar behavior has been seen in other phenomena and gives us a clue. Vialov and Zaretsky (1973) showed exponential creep deformation of clays. Gross (1947) showed that this exponential form is equivalent to a Voigt-type material model² that includes a very broad Gaussian range of parameters. Juarez-Badillo (1985) showed a Cole-Cole type deformation model that will be explored here to unite all of these observations.

The empirical relation developed by Juarez-Badillo is of the form

$$\epsilon(t) = \frac{\epsilon_{final}}{1 + (\tau/t)^d}. \quad (2)$$

Knowing that we need to find a power-law form to fit the observations, we can analyze this equation under the limit where the time of the experiment, t , is much less than the characteristic compaction time, τ , defined as when the sample has undergone exactly half of the final strain limit. This seems appropriate as we are making an effort to do lab experiments at much less than the time that we imagine these processing happening in the field. Equation 2 then becomes

$$\begin{aligned} \epsilon(t) &= \frac{\epsilon_{final}}{(\tau/t)^d}, \\ \epsilon(t) &= \frac{\epsilon_{final}}{\tau^d} t^d. \end{aligned}$$

We notice now that the strain at time 1

$$\epsilon(1) = \frac{\epsilon_{final}}{\tau^d}$$

and therefore

$$\epsilon(t) = \epsilon_1 t^d. \quad (3)$$

Not only does this equation fit well with the observed data, but considering only progressive quartiles of the data, constant and stable values for the regressed parameters ϵ_1 and d are obtained. This provides further justification in the selection of this model as this was one of the significant problems with use of the other models.

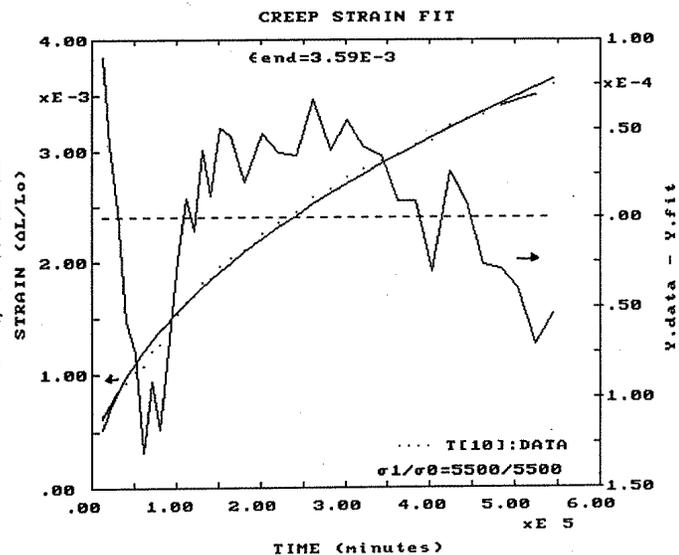
Now, assuming that our adoption of the Juarez-Badillo creep mechanism is correct, we have a model that helps explain our data. This fit implies several things:

²Yet another problem noticeable with these results is the apparent inconsistency of arriving at a Voigt-type model during a uniaxial strain test. Mavko (2001) indicates that the Voigt model is normally associated with iso-strain, uniaxial stress experiments.

- All tests are operating well in accordance with the assumption that $t \ll \tau$. Even the data shown in Figure 4 fit the exponential model nicely.
- Under these time constraints we cannot hope to solve for both ϵ_{final} and τ , but only for their quotient.
- Further research is needed to understand how a sample can achieve such self-select its mode of operation.

Figure 4: One year hold uniaxial creep test. Exponential function still fits meaning $t \ll \tau$ and implying that reservoir material must have a characteristic creep time on the order of decades.(Dudley and Myers, 1994)

brad1-year [NR]



TIME SCALING AS SAVIOR

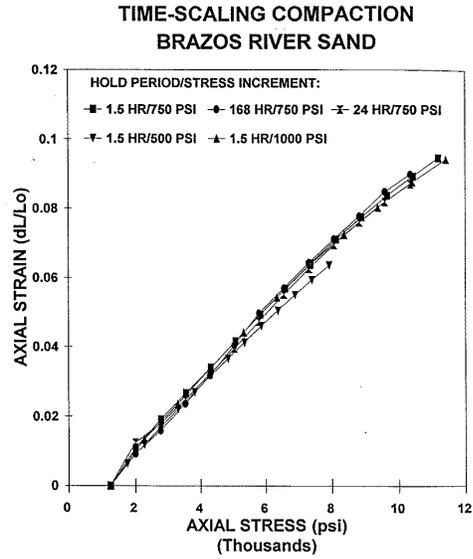
One of the initial interests mentioned earlier was the use of these experiments to determine the compressibility of reservoirs. As we noted above that it seems unlikely we will be able measure ϵ_{final} and τ , let us hope that we can at least accomplish this. The unconsolidated deep water reservoirs commonly found today can compress so significantly as to

- cause costly tubing collapse and demand pressure maintenance, and/or
- obviate water flood necessity for production efficiency.

These two seemingly contradictory results make it quite important to understand the validity of these tests in order to make correct decisions.

As was shown in Figure 3, that the final strain value of each test cycle was identical and independent of the hold time, we see that each test results in the same stress-strain behavior in Figure 5. We are now free to calculate the compaction curves as a function of pressure.³

Figure 5: Time scaling nature of the normalized hold-cycles results in each sample attaining the same strain value at the end of the hold time. Plotting final strain values from all cycles tested shows overlay of stress-strain behavior. Also important in this testing methodology is the pre-stressing of samples to pseudo-depth pressures. This removes 'closure' errors from consideration of the results. (Dudley and Myers, 1994) brad1-ss [NR]



This result will tell us the expected uniaxial compaction of a reservoir during the drainage phase that increases the vertical effective stress (VES) as pore pressure draws down ($VES = \sigma_1 - P_{pore}$), under the caveat that the production time does not approach the characteristic time τ . If we violate this assumption, we will over-predict the amount of compaction the reservoir will experience.⁴ This allows us to present plots of the regressed parameter ϵ_1 versus pressure such as Figure 6.

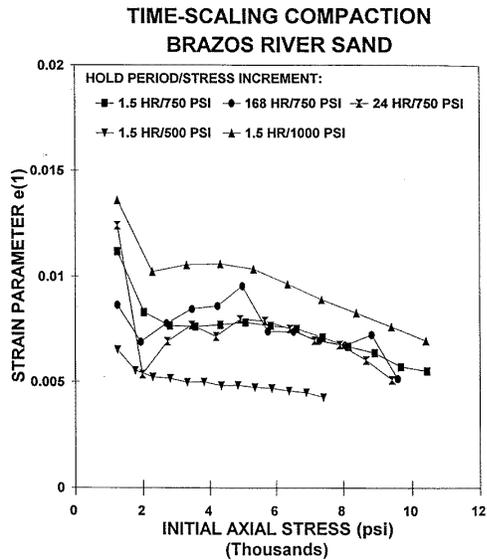


Figure 6: Parameter $\epsilon(1)$ from all tests. All 750 psi steps overlay. (Dudley and Myers, 1994) brad1-e1 [NR]

³We must remember that this is not an elastic phenomenon, and no work has been done thus far to determine to what extent there may or may not be hysteretic effects when attempting to reverse this process. This concept could be especially important for aquifer studies and late-stage pressure maintenance projects.

⁴It is also good to remember for the porosity loss due to compaction that the pore volume compaction is related to the uniaxial compaction by $C_p = C_z/\phi$.

Data from several pressure step magnitudes are presented. By dividing by the magnitude of the pressure step, we effectively calculate C_z from equation 1 and watch the curves collapse onto one another.

Figure 7: Remarkably, uniaxial compaction coefficient C_z as calculated by 5 different testing methodologies remain consistent. (Dudley and Myers, 1994) `brad1-cz` [NR]

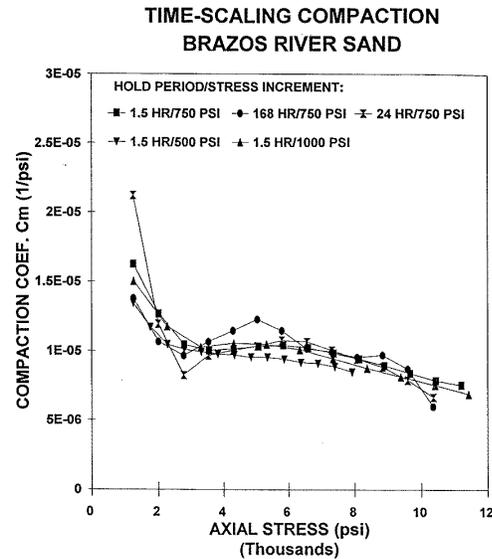


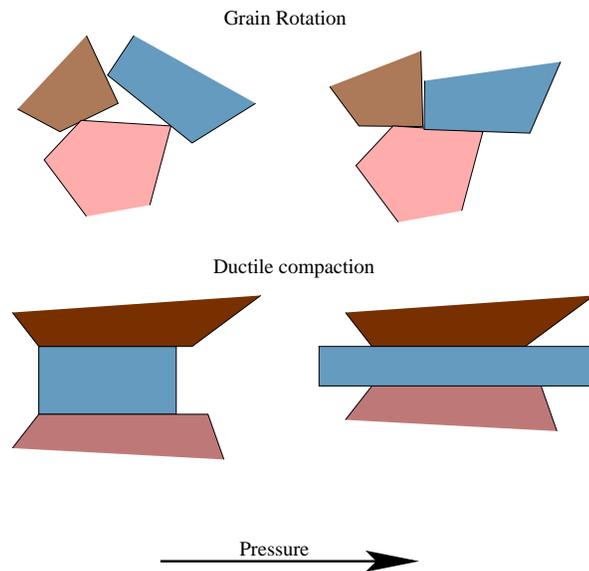
Figure 7 shows the quite compact distribution of the C_z parameter for 5 different samples each calculated with different stress-step and hold-time combinations. Thus the time scaling nature of these tests results in the pressure scaling of the tests as well due to the compulsion that the different tests must satisfy identical stress-strain behavior (as shown in Figure 5).

MECHANISMS

So far, sufficient evidence has been accumulated to comfortably apply laboratory compaction tests of any convenient hold time ⁵ to faithfully characterize the reservoir. However, it is unclear to me as to physically how this seemingly *double-slit* style interaction manifests itself in the changing rocks. There seem to be very few available mechanisms for accommodating creep behavior. Shown in Figure 8 are cartoons illustrating the three available to the rock without grain crushing. The problem then is how, through these limited mechanisms, can the rock realize the necessary broad distribution of characteristic relaxation times? It seems reasonable that both mechanisms may enjoy some breadth in time-scale simply as function of heterogeneity of mineralogy and sorting. To truly achieve the smooth single mode distribution required however implies that the ductile grains act as impulse sensitive springs and the grain rotation is almost fractal. These requirements seem difficult enough to achieve without the further requirement that they overlay and interact nicely together. The experiments shown here are all done on manufactured plugs of frozen loose river sand described as sub-angular, moderately well sorted, feldspathic and fine grained. Similar tests on Gulf turbidite reservoir core samples

⁵Convenient does not necessarily mean shortest. Notice that if a hold time of hours is left or forgotten for a longer time period, the sample will begin to re-scale its creep mechanism toward the longer option. Weird? Yes.

Figure 8: Three compaction, without crushing mechanisms: 1) grain rotation, 2) ductile compression of grains, or 3) compression of pressured pore space (the transition of an over-pressured state to a grain supported frame). [bradl-mech](#) [NR]



show similar results. It is my suggestion then to perform a similar suite of tests on manufactured grain packs in an effort to understand how the various mechanisms are responsible for which parts of the scaling phenomenon. To this end samples of various sorting, grain size, grain shape, mineralogy, and combinations of all of the above could be manufactured and tested under uniaxial strain tests. A further help in identifying the processes involved would be to take intermediate images of the sample to show the nature and state of the undergone transformations similar to Cadoret's partial saturation experiments (from Mavko (2001)).

Further, once these processes and their limiting and defining characteristics are well understood, the development of type curves for various material would be very valuable.

IMPLICATIONS TO 4D

Accepting the previously outlined rock physics conundrum, we can now think about the implications of this creep behavior, and our ability to capture it in the lab, to the repeat seismic experiment. As most of the rock physics field has historically been dedicated to the application of acoustic manipulations and pressure dependencies there is much work available concerning the elastic sensitivities of rock, packs, and beads to various pressure regimes (Mavko, 2001). All of this understanding however hinges on the basic immutability of the mineral frame (elastic deformation) until a yield point from which a sample cannot recover (fracture or grain crushing). We begin to realize that our discussion thus far begins to impinge on the underlying assumption for the classic rock physics justifications for 4D seismic: namely that the rock frame does not change. The previous sections show that it does. Thus there are no constants to hold on to for any experiment, be they fluid substitution or seismic. We might as well be dealing with an entirely new reservoir. To add to the conclusions to the creep investigation section above, it is imperative to perform experiments to not only understand the creep mechanisms, but also their effects on sonic velocities or acoustic moduli. It is obvious that the

Figure 9: Experiment illustrating the effective pressure law by increasing and decreasing effective pressure while keeping the pore pressure constant. Minor hysteresis observable, but obviously an elastic phenomenon. (Mavko, 2001) brad1-pres [NR]

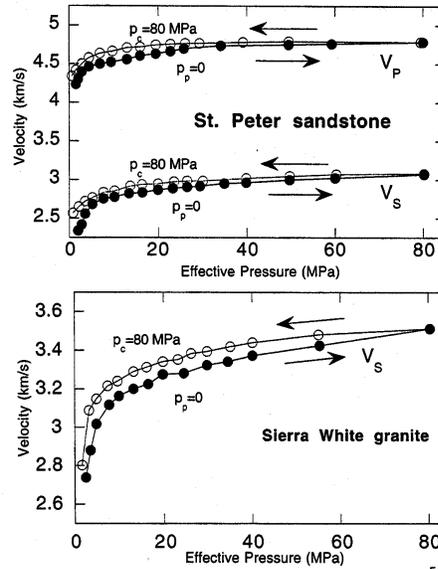


Figure 10: Different micro-distributions of pore fluids and gas within the porosity of the rock frame significantly effect the velocity dependence on water saturation. (Mavko, 2001) brad1-sat [NR]

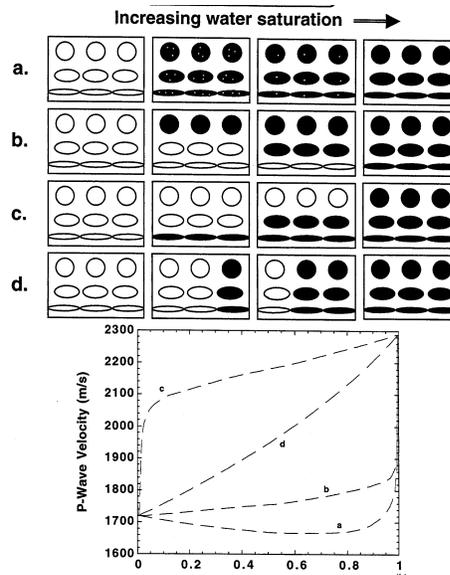
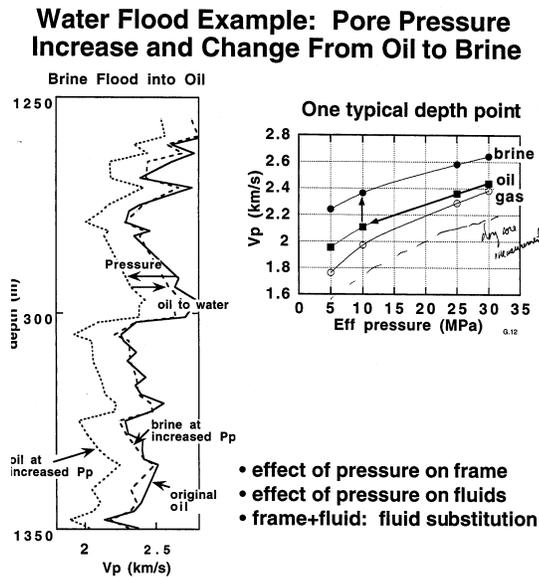


Figure 11: If fluids and elastic pressure dependence are understood, the seismic evolution of a reservoir can easily be modeled with a few simple plots. Fluid substitution calculated in this instance with Gassman model from dry lab data. (Mavko, 2001) `brad1-model` [NR]



various transformations depicted in Figure 8 will result in new elastic moduli as we are basically manufacturing a new rock sample. This is quite troubling when the second conclusion of Juarez-Badillo creep behavior is remembered. Solving for ϵ_{final}/τ rather than for the two independently removes our ability to predict end-state compaction and the time (supposed to be analogous to production life of a reservoir) to achieve it. While this seemed a disappointing but not insurmountable obstacle previously, it could be a real show stopper in the 4D context. To that end, the experiments mentioned previously need be undertaken with the addendum of velocity measurements along the way. It is very possible that while the samples can show time-scaling of creep mechanisms, parts of the distribution of mechanisms may indeed exhibit very different manifestations in elastic parameters.

REFERENCES

- Dudley, J., and Myers, M., 1994, Measuring compaction and compressibilities in unconsolidated reservoir materials via time-scaling creep: Measuring compaction and compressibilities in unconsolidated reservoir materials via time-scaling creep:, Rock mechanics in petroleum engineering, 45–54.
- Gross, B., 1947, On creep and relaxation: Journal of Applied Physics, **18**, 212–221.
- Juarez-Badillo, E., 1985, General time volume change equation for soils: General time volume change equation for soils:, Eleventh international conference on soil mechanics and foundation engineering, 519–530.
- Mavko, G., Rock physics:, Course notes, Spring quarter 2001.
- Segall, P., 1998, A note on induced stress changes in hydrocarbon and geothermal reservoirs: Tectonophysics, **289**, 117–128.

Vialov, S., and Zaretsky, Y., 1973, Kinetics of structural deformations and failure of clays
Kinetics of structural deformations and failure of clays, International conference on soil
mechanics and foundataion engineering, 459–464.

Zoback, M., Reservoir geomechanics:, Course notes, Fall quarter 2001.

Short Note

AVA attributes (2 for 1 special)

Brad Artman¹

INTRODUCTION

Once satisfied that an amplitude, as produced by the user's favorite migration process/algorithm, contains sufficiently accurate and interesting information, the use of that knowledge must be simple and enlightening. Herein I will introduce an AVA (or O if you insist) attribute that is easily calculated, quickly interpreted, and physically meaningful. This attribute will be derived from a different, though related, cross-plot that shows significant improvement over the classic *Intercept-Gradient* plane.

While it is clear that there are likely few geophysicists truly excited about the introduction of yet another seismic attribute, I believe that Shale-Normal Amplitude (SNA) analysis described herein can greatly ease and quantify the interpretation of AVA data. The premise is very simple and the output concise and enlightening. The attribute is derived from a *Compression-PseudoShear* reflection coefficient plane that allows significantly more insight into the lithologies at depth than the traditional *Intercept-Gradient* ($A - B$) plane.

THE PSEUDO-SHEAR REFLECTION COEFFICIENT

Most practitioners are by now familiar with $A - B$ plots for AVA analysis. The SNA attribute operates on a permutation of the $A - B$ plane described below. The Zoeppritz equation approximation for small physical contrasts over intermediate angles presented by Shuey (1985) is:

$$R(\theta) \approx A + B \sin^2 \theta, \quad (1)$$

where

$$A = \frac{\Delta\rho}{2\rho + \Delta\rho} + \frac{\Delta V_p}{2V_p + \Delta V_p} \quad (2)$$

and

$$B = A - 2C,$$

¹email: brad@sep.stanford.edu

with C having the form

$$C = \frac{1 + \eta}{2} \frac{\Delta\rho}{2\rho + \Delta\rho} + \eta \frac{\Delta V_s}{2V_s + \Delta V_s} \quad (3)$$

where

$$\eta = 4 \frac{V_s^2}{V_p^2}.$$

Δ quantities are the differences layer 1 minus layer 2. Density and velocities (without Δ) are those of layer 2. Also, I have reordered the usual relationship to lead the reader toward the goal of orthogonalizing the AVA plane between compressional and shear axes. With the above development of the Shuey equation, we see A is the normal incidence compressional reflection amplitude and B contains both normal incidence and angular dependence. Other authors, such as Castagna et al. (1998), have attempted to simplify this relationship by expressing B as a complicated function of A multiplied by new (empirical) fitting coefficients. While facilitating ever more cross-plotting possibilities, axes remain mixtures of compressional and shear quantities and yield little more insight.

However, with the above formulation of Shuey equation (1) we can separate B into the compressional, A , and *pseudo*-shear, C , reflection amplitude coefficients. Notice the very parallel structure of the *pseudo*-shear, equation 3, to the normal incidence term, equation 2, and its independence from compressional velocity contrast. Encouragingly, the *pseudo*-shear expression contracts to the normal incidence shear reflection coefficient when the compressional to shear velocity ratio γ equals 2.

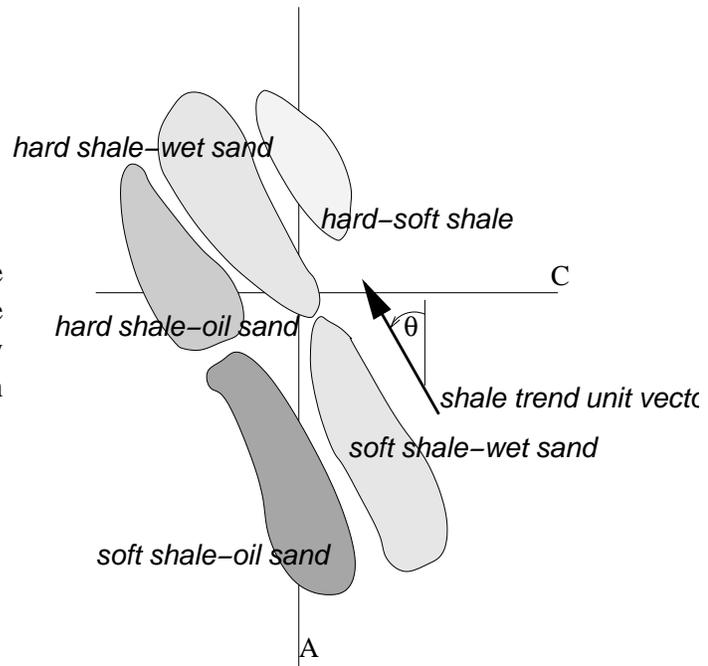
This separation (effectively between $A(\Delta V_p)$ and $C(\Delta V_s)$) will help guide our intuition by isolating the seismic reflection amplitude into two components that have meaning in an AVA sense. Rather than plotting on the $A - B$ plane, we can utilize the $A - C$ plane and avoid an ordinate with codependency of compressional and shear velocity boundary contrasts.

With the insight gained from the definitions of A and C , we can now intuitively understand that trends in this AVA plane are due to the variation of real rock properties. **Figure 1** indicates the relative position of simple targets in this space. Because I have defined density and velocity as the properties of the lower layer, the negative values of A show a change from harder to softer intervals and the opposite is true for the positive values. Therefore we can understand something of the nature of the bounding layers of an interval as harder bounding lithologies will trend to more negative values of A .

We also know that the shear velocity of a porous medium increases as we lower the density of the included fluid. This tells us that the ΔV_s will be negative as we consider a water filled medium versus a gas filled one and this will increasingly drive the value of C more negative. The interplay between these compressional and shear forces results in the normal NW-SE trend of the data cloud in **Figure 1** indicating hard bounding rocks upward and soft ones downward.

Gratwick (2001b) outlined the promise and difficulty of prospecting AVA anomalies on an $A - B$ plane as explained by Castagna and Swan (1997). Both authors stress the importance of the distance away from the background trend for the analysis of a prospective event. Attempting to quantify this, Gratwick (2001b) calculates the product of A with B , then masks the cen-

Figure 1: With the clarity of structure of equation (3) for C , simple test case plots can be readily manufactured by applying it to mental scenarios within this cartoon. brad2-cartoon [NR]



ter mass of reflection amplitudes that are assumed to be background values (non-prospective shale-wet sand or shale-shale reflections). This process is shown in **Figure 2**. The flaw in this method is the dull spoon that differentiates anomalies from background. Not only is the scalpel dull, but this methodology only appreciates a single model type. More practically, the clumsy transfer in and out of SEP architecture for graphical definition of the mute zone can dissuade all but the most committed from utilizing this tool.

Figure 3 (i) shows the standard Slope-Intercept AVA plot, while (ii) shows the transform to the Compressional-PseudoShear plane. The data are generated from a synthetic provided by BP and explained in detail by Gratwick (2001b). While immediately displeasing, these two plots will highlight the power of the $A-C$ plane when inspected. First note the strong zero presence on the intercept-axis of the $A-B$ panel. This is modeled data, boring, and makes our unit vector for the shale trend very simple ($\theta = 0$). In an attempt to provide a small measure of believable scatter (make this plot less boring), a bandpass filter was run over the AVA attributes.² This contributes to a few bothersome artifacts, but are easy to neglect. These include: data present to the left of shale trend (bandpassing 0 returns negative values), diagonal sub-trends of events, and incomplete orthogonalization. We see that due to the presence of both shear and compressional velocity contrasts in the formulation of B , the transition of reflections on the $A-B$ plot from water to oil to gas takes place along a line with an acute angle to the background trend. This leads to one of the paramount problems with interpreting AVA anomalies as explained in Castagna (1997). The $A-C$ plane, enjoying an ordinate quantity that is a function only of a change across the boundary of the shear velocity, shows nice perpendicular³ departure from the axis of the compressional reflection coefficient.

²Without such, only four points are seen on the plane.

³I really think it would be without the bandpass issues

Figure 2: A vs. B scatter-plot with mute fairway defined. Gratwick (2001b) `brad2-plot2` [NR]

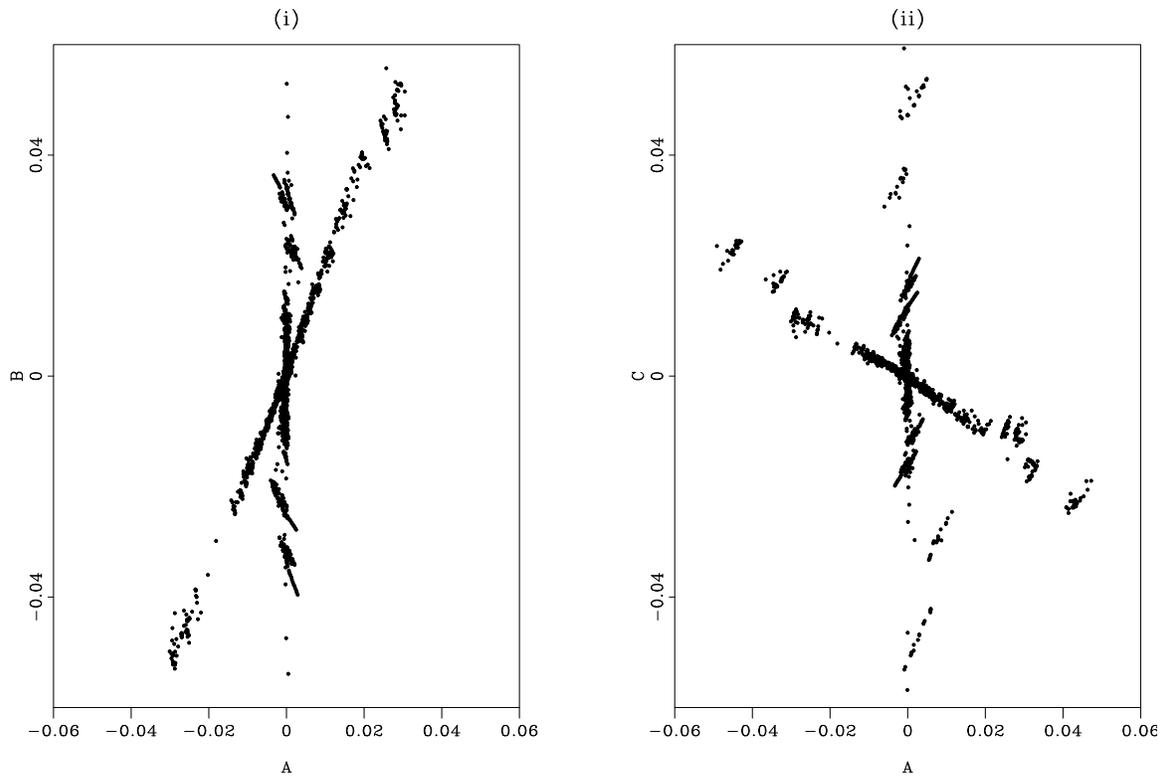
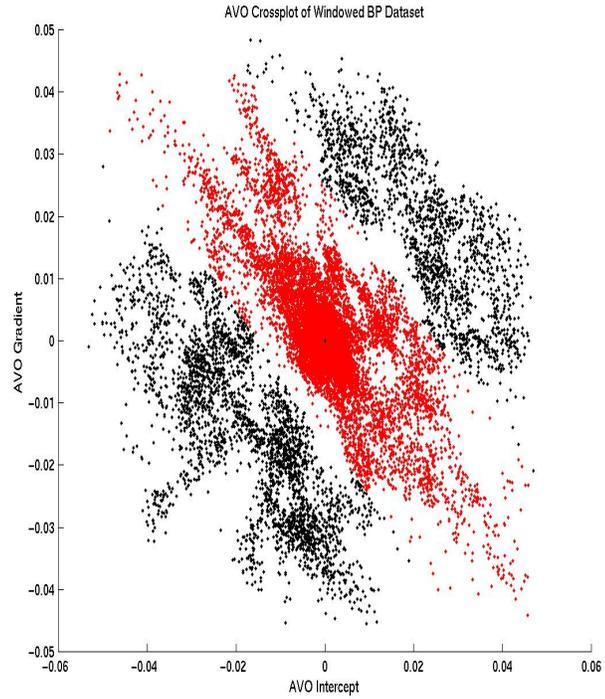


Figure 3: The ordinate axis is transformed from B , Shuey's gradient, to compressional the pseudo-shear reflection coefficient C . `brad2-planes` [ER,M]

SHALE-TREND NORMAL AMPLITUDE

Intuitively now, we have developed an understanding for where on the $A - C$ plane to expect prospective events and a little about what a location on the plane indicates about the rocks. The perpendicular distance away from the shale trend explained above is what truly quantifies an AVA anomaly. Therefore, if we can define a unit vector that accurately describes the shale trend, the cross-product with a vector containing the values of A and C for any point in the subsurface will interpret whether that point is *anomalous* or not. This quantity I will call the Shale-normal Amplitude. Simply stated, defining the reflection coefficient vector

$$\vec{R\bar{F}C} = (A, C)$$

and the shale trend background vector

$$\vec{S} = (\cos\theta, \sin\theta)$$

the Shale-Normal Amplitude is

$$SNA = \frac{|\vec{R\bar{F}C} \times \vec{S}|}{|\vec{S}|} = \frac{A\sin\theta - C\cos\theta}{|\vec{S}|}. \quad (4)$$

USE AND APPLICATION

Piggy-backing directly on the development and results of (Gratwick, 2001b), and assuming that all necessary precautions in dealing with amplitudes as a function of angle have been taken into account (as detailed in (Gratwick, 2001a)), we can compare the $A * B$ mute method with the SNA analysis. Analogous with the mute method, a threshold value for the SNA will need to be selected.⁴

This value can be chosen by inspection of a SNA histogram. We can hope for clean separation of roughly Gaussian distributions of shale-wet sand reflections and shale-oil sand (and/or shale-gas sand) resulting in a bi- or tri-modal distribution of reflections. As discussed above, the less dense the pore-filling fluid, the higher the SNA value. Therefore, assuming only that most intervals are water-bearing, we can confidently pick the first mode as indicative of water reflections. **Figure 4** shows an ideal distribution of SNA, while **Figure 5** shows the histogram for the synthetic data.

An even more powerful approach is to extract the SNA attribute along an event and plot it as a function of depth or time. As the event crosses fluid levels, we will see nice step-changes in the SNA value.⁵

⁴This method is not salt-proof. Like so many other things salt really screws up things. All analysis here-in has windowed out the region of the salt body.

⁵While these well behaved model amplitudes are picked by peak amplitude, the author will freely rant upon request about the importance of using some product of amplitude and thickness of wavelet when analysing real data.

Figure 4: Ideal distributions of SNA attribute. Threshold values for color tables or clipping can be set with these in mind for fast and furious prospecting or horizon mapping and volumetrics. `brad2-ideal` [NR]

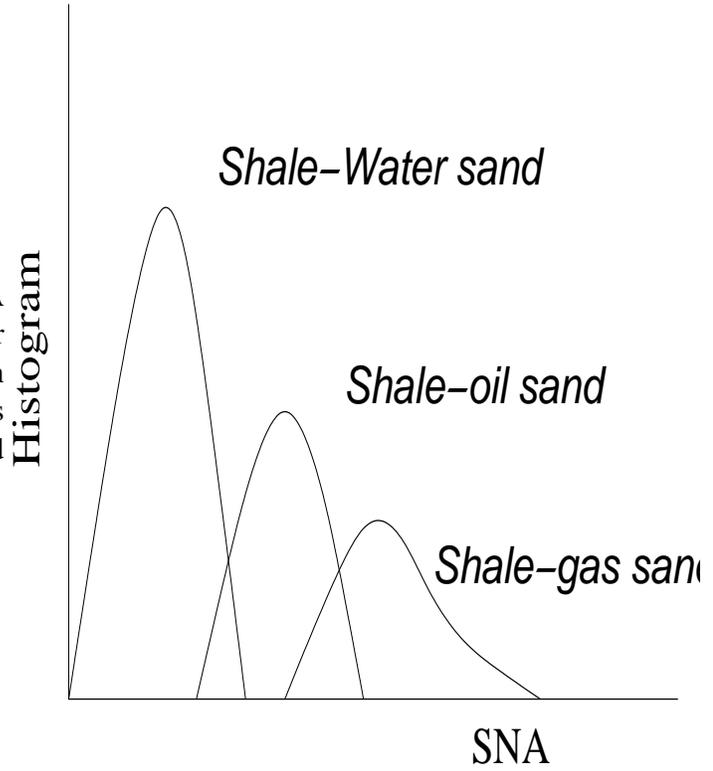
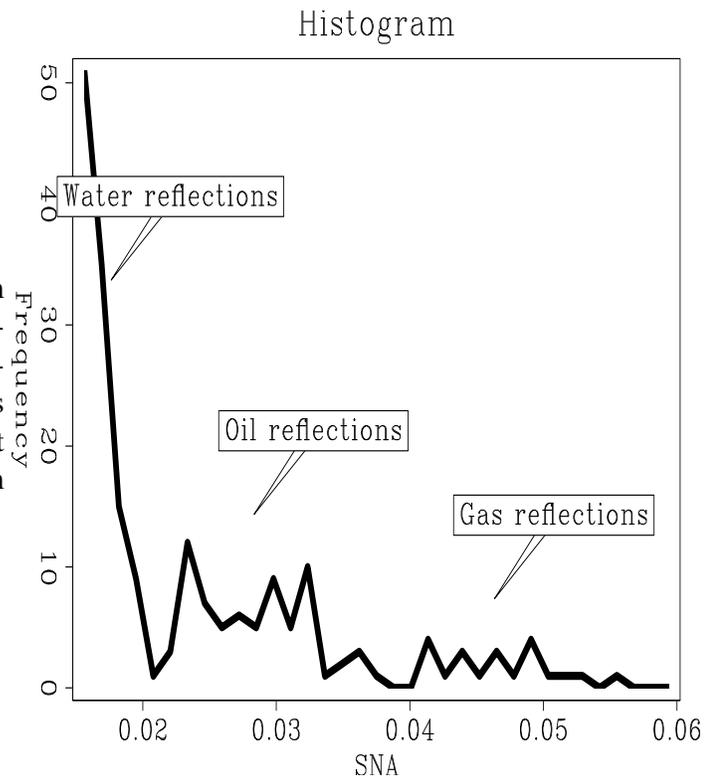


Figure 5: Histogram of SNA from modeled then migrated from the density/velocity grid shown in 6. The extreme preponderance of zero values in the model makes the plot ugly, but the tri-modal nature of the histogram is evident. `brad2-histo` [ER]



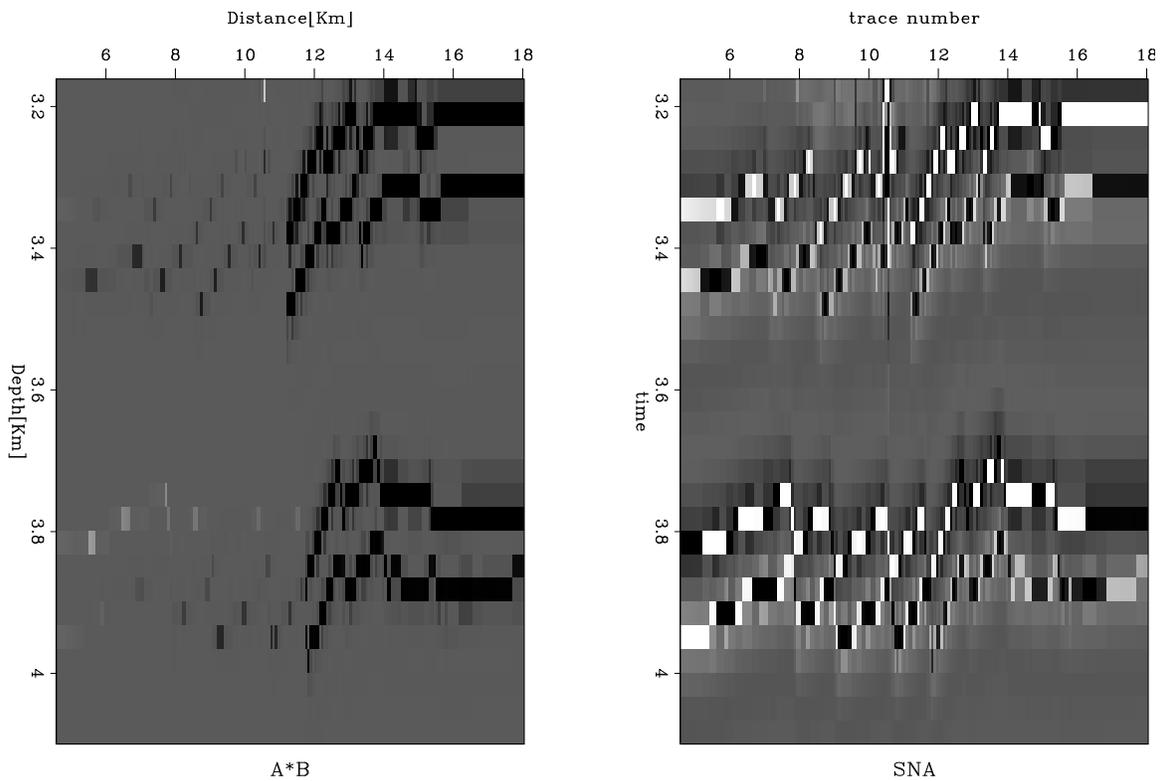
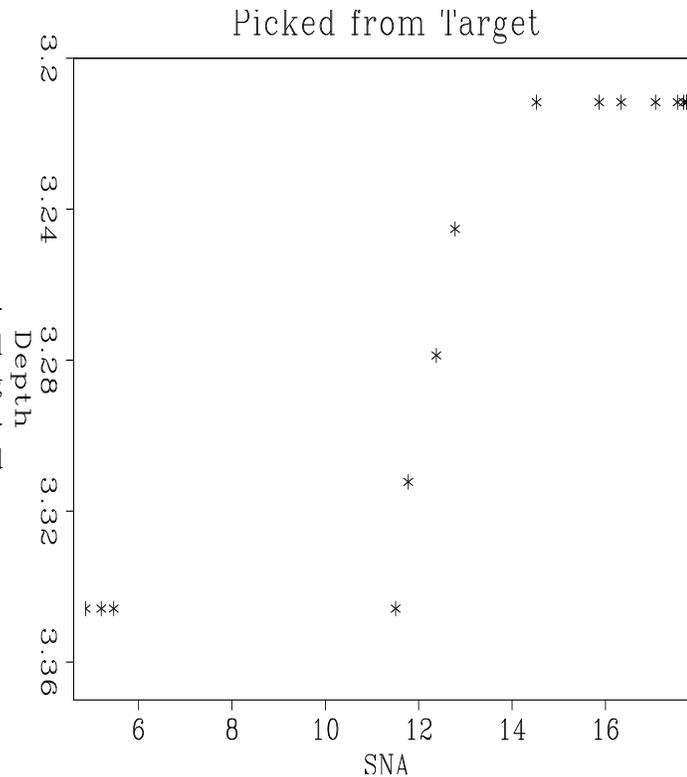


Figure 6: Zoom of “reservoirs” from V_p , V_s , and ρ models. Gratwick (2001b) shows the entire model and details the processing steps. White denotes high A*B value (down-dip water reflections have already been muted). Notice the accuracy of the SNA model showing prospective AVA character on the top reservoir reflection, but not the bottom reflection. **Figure 5** shows picks along the upper-most reflector. [brad2-models](#) [ER,M]

Figure 7: While fault blocks complicate the plot somewhat, fluid level remains obvious, and a prospecting threshold of 12-14 is comfortably assigned. Values were not normalized by $|\vec{S}|$. `brad2-event` [ER]



To use this type of analysis on well or seismic data (integration!), one is only required to perform the trivial calculations of equations (1) and/or (3) to generate the $A-C$ plane. Further, after estimating the Shale trend angle on the plot,⁶ the SNA attribute can be calculated and used as a prospecting or reservoir characterization tool.

The next step in proving the value of this concept will be to apply these simple transformations to single events known to have fluid contacts. Such a data set is thought to exist in-house, and a thorough description of our success/failure with this analysis will be included in the next report.

ACKNOWLEDGMENTS

I would like to acknowledge Doug Gratwick's work in SEP108 upon which the examples used herein are built.

⁶It is unclear whether simple regression on this plane could result in incorrect unit vectors without the dismissal of anomalous points by an interpreter. What is encouraging is that in the test thus far, these calculations are very insensitive to the precision of the θ .

REFERENCES

- Castagna, J. P., and Swan, H. W., 1997, Principles of AVO crossplotting: AAPG Mid-Continent Section Meeting; Abstracts; AAPG Bulletin, **81**, no. 8, 1348.
- Castagna, J. P., Swan, H. W., and Foster, D. J., 1998, Framework for AVO gradient and intercept interpretation: Geophysics, **63**, no. 3, 948–956.
- Gratwick, D., 2001a, The accuracy of wave-equation migration amplitudes: SEP-**108**, 63–72.
- Gratwick, D., 2001b, Amplitude analysis in the angle domain: SEP-**108**, 45–62.
- Shuey, R. T., 1985, A simplification of the Zoeppritz equations: Geophysics, **50**, no. 4, 609–614.

Short Note

Ricksep: Interactive display of multi-dimensional data

Robert G. Clapp¹

INTRODUCTION

SEP has always been interested in displaying and interacting with multi-dimensional datasets (Claerbout, 1981; Sword, 1981; Ottolini, 1983; Ottolini and Rocca, 1986; Ottolini, 1988, 1990; Biondi and van Trier, 1993; Clapp and Biondi, 1994; Clapp, 1995; Mora et al., 1995, 1996; Clapp et al., 1994) In 1997 Rick Ottolini provided SEP with an extension of his old X11 movie program written while at Unocal. In the last two years I have added several additional features to the code. These changes include:

picking Expansion of the picking ability including some degree of auto-picking.

annotating The ability to annotate figures and then output the same view using SEPlib programs.

velocity The ability to do velocity analysis.

multiple depth The ability to work on X environments that do not have support pseudo-color.

multiple views The ability to have multiple views of the same-dataset or two equivalent datasets.

This paper explains how to use each of these new features.

PICKING

The ability to pick points in a multi-dimensional cube was part of the original `Rickmovie` program. The original design was limited in the respect that picks could not be differentiated. Therefore picking multiple surfaces required either erasing each individual pick or restarting. To allow multiple objects to be picked I added a marker component to the stored picking information. The marker is a single text character that can be changed by bringing up the picking

¹email: bob@sep.stanford.edu

menu (hold down ‘c’ and press the right mouse button). The `TEXT` field of the picking menu is used to set the active marker. Once the `Change` button on the picking menu is hit each new pick will be displayed with the new marker. Figure 1 demonstrates the concept. One object is marked with ‘b’ the other with ‘a’. When the picks are written out, the mark character will also be written.

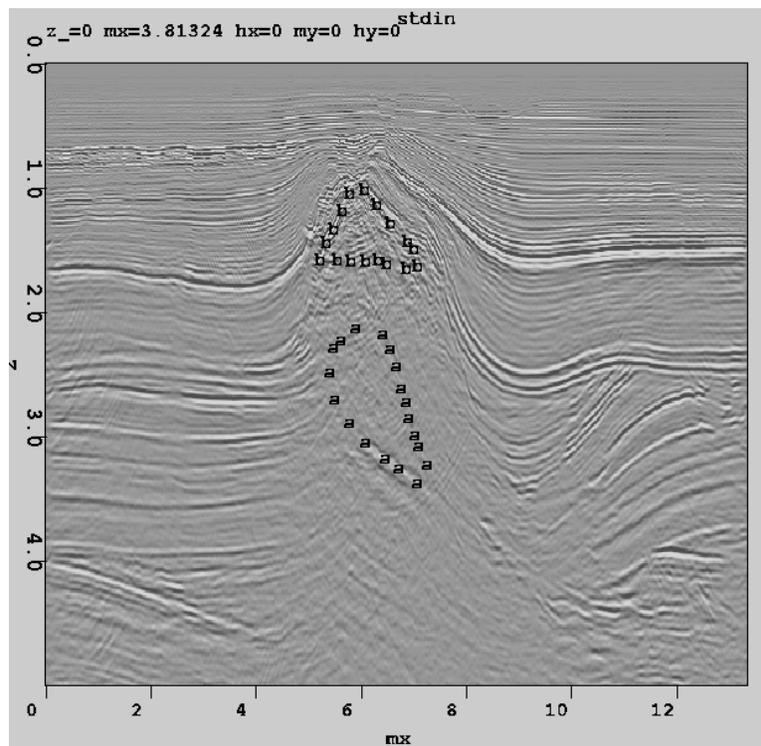


Figure 1: An example picking multiple surfaces. `bob2-picking` [NR]

Another limitation of the original design was the inability to do auto picking. I added a couple rudimentary auto-picking features. The first is the ability to ‘snap’ picks to similar locations (activated by holding down ‘b’ and the right mouse button). To figure out the ‘snapped’ location all the points within the currently viewed cube, with the active marker are selected. A super trace is built by taking a region (around a user-specified dependent axis) the selected point and summing. Cross correlation is then run to find the shift at each location that produces the highest correlation with the super-trace. A new super trace is then constructed from the new point locations. The process is repeated five times. The dependant axis and the search radius are both changeable through the picking menu.

In addition to the ability to snap points a simple auto picker is included. The auto-picker uses a basic region growing algorithm. It uses the simple dip and coherency calculation method of Claerbout (1992). Figure 2 demonstrates its use. In this case three points were seeded for the algorithm.

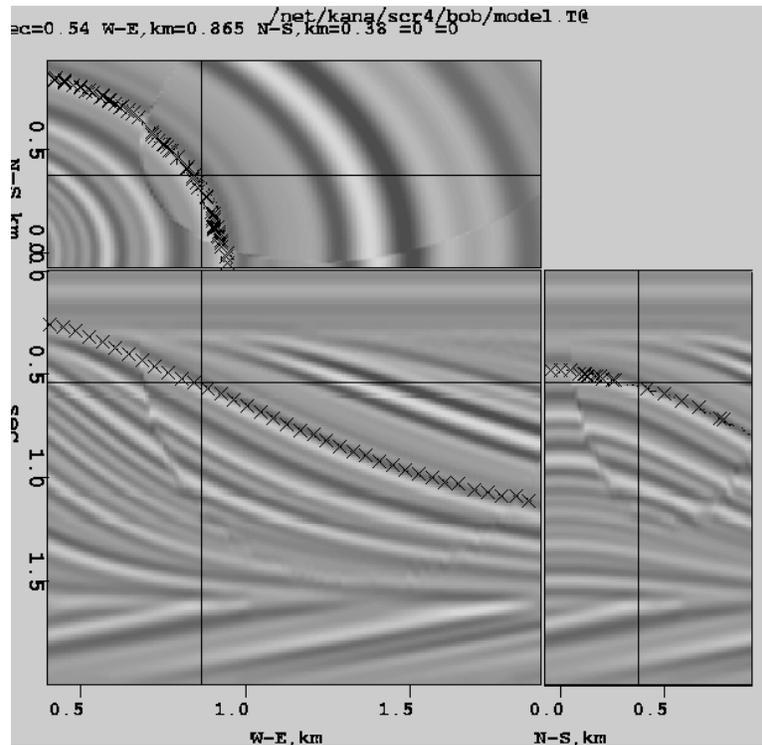


Figure 2: The auto-picker in Ricksep. The 'x' were seeded with 3 points. `bob2-auto` [NR]

ANNOTATION

Figure annotation is often necessary to direct a reader/viewer's eye, especially in static media such as paper. SEP's main graphics package, `vplot` (Cole and Dellinger, 1989; Dellinger, 1989), lacks all but rudimentary interactive ability. Rather than invest the time required to rewrite what is generally a bug-free, effective package, I decided to add some basic annotative ability to Ricksep. Figure 3 shows an example of the different types of objects that can be used for annotation. To do annotation the `Annotate` button must be selected from the menu. Many of the zoom and picking functions are then replaced with annotation functions. The table below shows the various options.

Keystroke/Mouse	Function
Left-button, drag	Draw an oval
Center-button, drag	Draw a rectangle
Center-button, click	Add text
a+Right click	Add point for polyline
f+Right click	Finish polyline object

Ricksep is designed to be an interactive viewer. It generally can't compete with `vplot` for static displays. In order to take advantage of the abilities of both packages I added the ability to output the SEPlib commands to make an equivalent `vplot` figure. Figure 4 shows the result

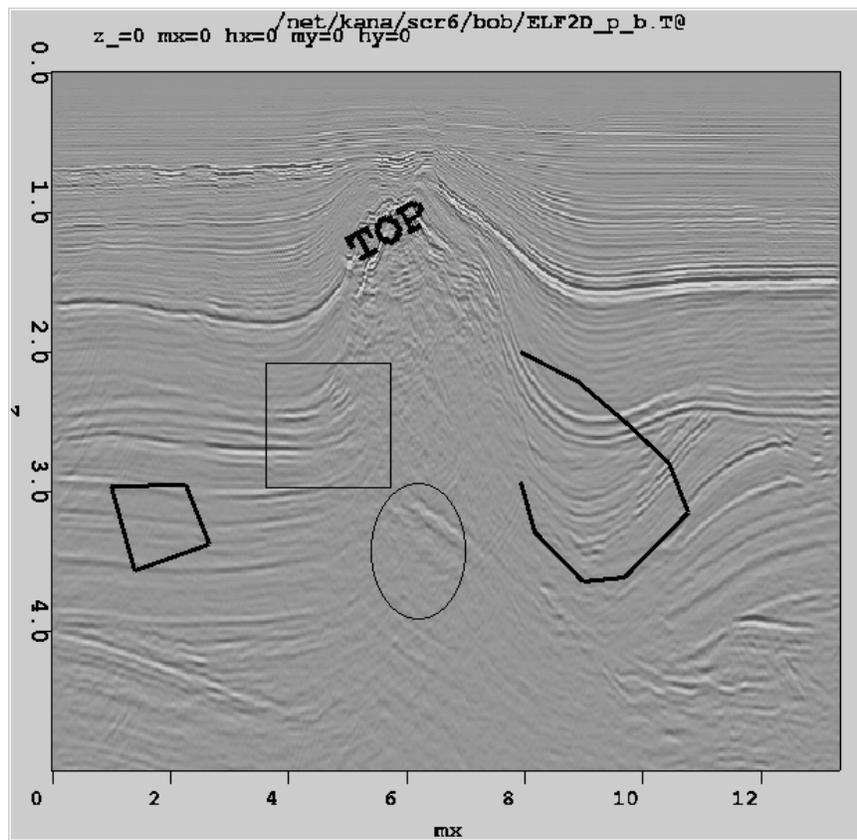


Figure 3: A view of Ricksep of a 2-D slice annotated with a box, text, an oval, and two poly-lines. `bob2-annotate` [NR]

of running the commands outputted with the `section->display->output` commands menu option.

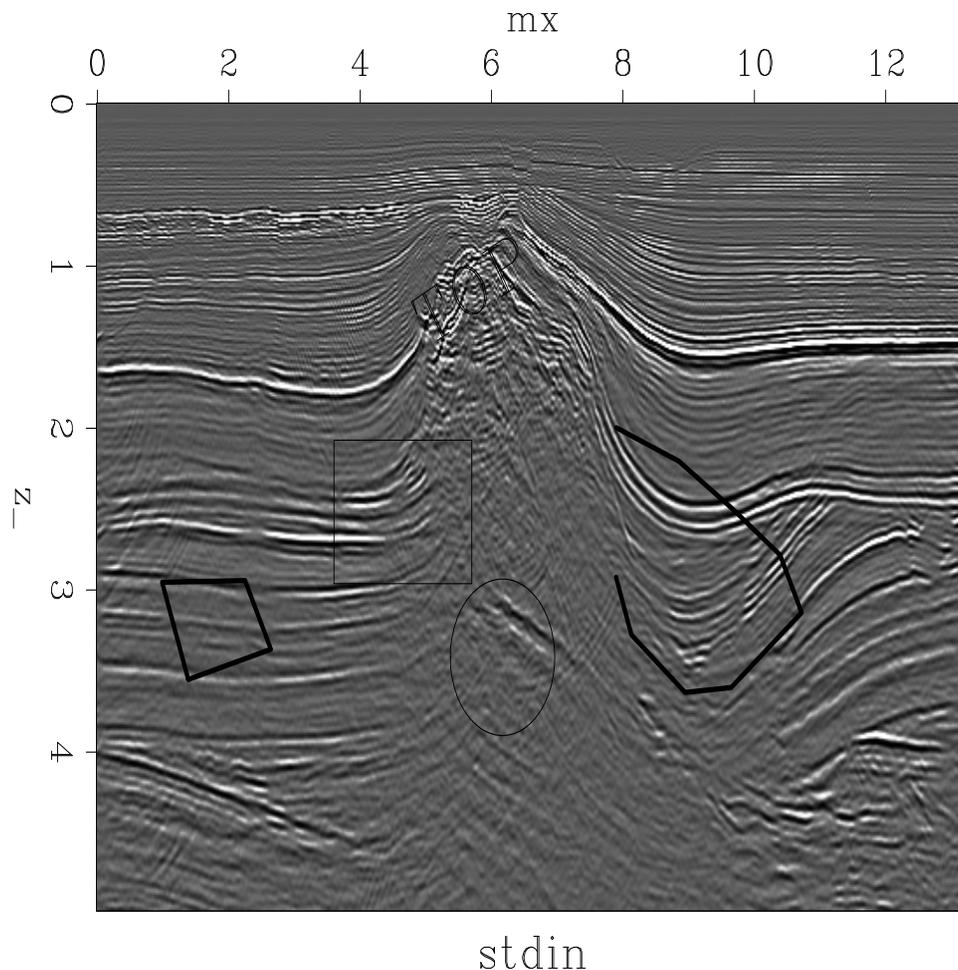


Figure 4: A 2-D slice annotated with a box, text, an oval, and two polylines, produced using the `output` commands option of Ricksep. `bob2-ano-v` [NR]

MULTIPLE VIEWS

The original Rickmovie was very effective in displaying 3-D cubes, but it wasn't good at handling four and five dimensional datasets. Viewing dimensions larger than three is generally a challenging problem. Rather than create a set viewing method I instead added the ability to have multiple, linked, views within one Ricksep session. Each view can contain the same or a different, equivalent sized dataset, and can correspond to any possible combination of up to five axes. As expected adding this level of flexibility quickly causes an explosion in the number of parameters to the program, and leads to what amounts to the development of a scripting language. To allow easy uses of these features I created three simple perl scripts for some of the most common display possibilities.

Displaying 4 and 5-D datasets

The first of these scripts shows one way of displaying four and five dimensional datasets. In this case the script 'Rick4d' is called. The resulting view, Figure 5 shows the first, third, and fourth axes in the left display and the first, and second axis in the right display. So in this case we are looking at CRP gathers in the right display at the location specified in the left display. For five dimensional data the right display becomes axes 1, 2, and 5.

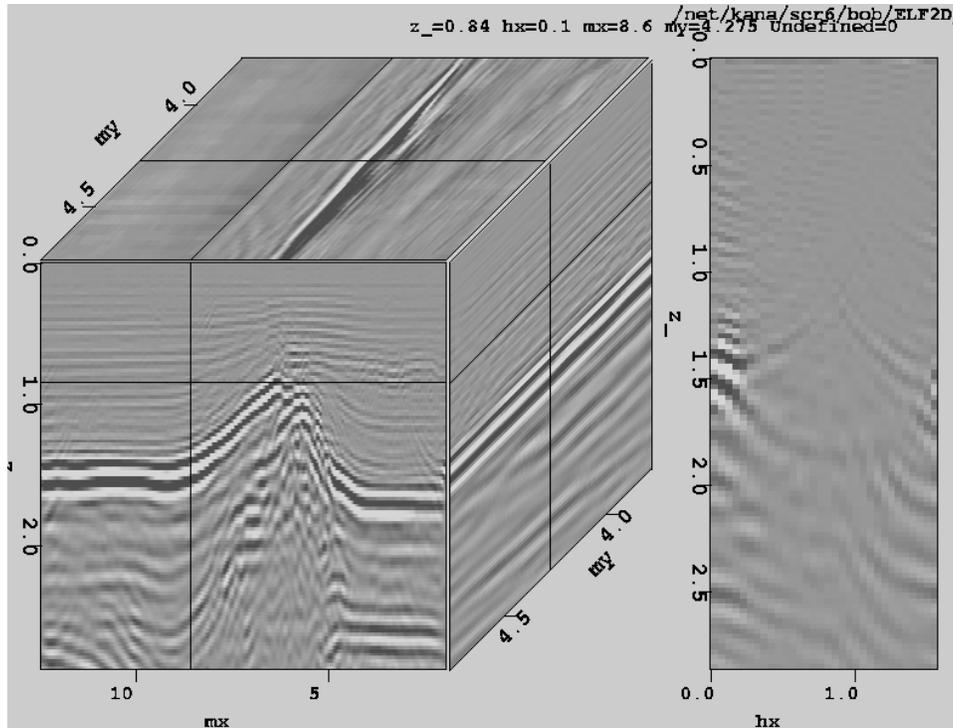


Figure 5: Viewing of a 4-D dataset in Ricksep. The left panel shows CMP locations, the right panel is a migrated CRP gather. The two panels are linked so that the CRP gather corresponds to the midpoint specified by the cross-hairs in the left panel. `bob2-multid` [NR]

Comparing datasets

As long as the datasets are consistent the different views can contain different data. Figure 6 shows two different datasets using the `Rick_multi` script. The views are again linked, but each dataset has its own pickset, display format option, etc.

Velocity Analysis

The third variation is accessed through the `Rickvel` script. It displays CMP gathers in the left display, a semblance panel at the given CMP location, and the NMO corrected gather

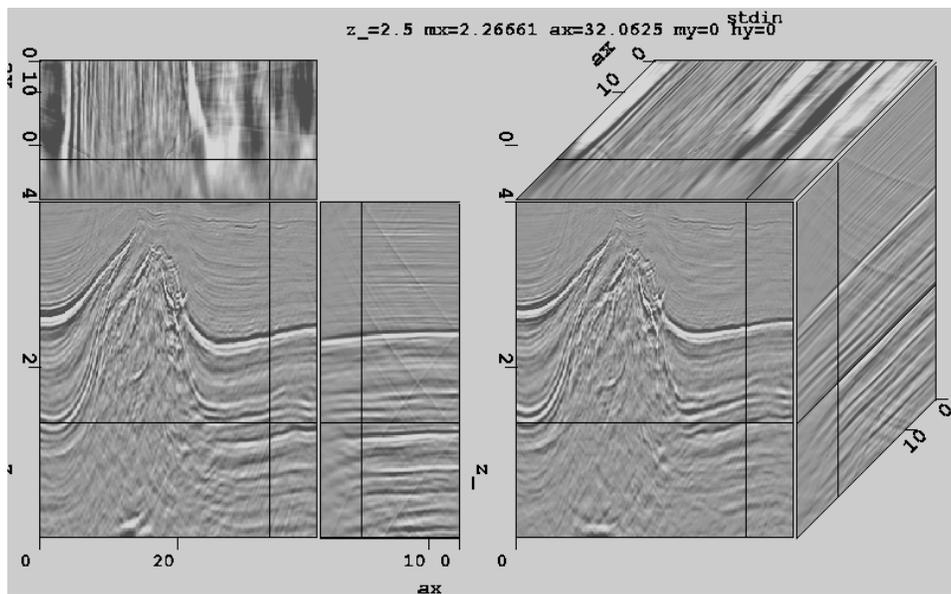


Figure 6: Comparing multiple datasets. The two panels are two different migration results. Moving locations in one gather causes the same position to come up in the second gather.
bob2-multiple [NR]

(Figure 7). It lets you pick the velocity function on the semblance panel and automatically updates the NMO panel. When you quit Rickvel, it outputs the RMS velocity function.

MULTIPLE DEPTHS

`Rickmovie` relies on the pseudo-color X11 Visual. This visual is the default on most Unix systems. On most current Unix platforms several visuals are supported. Unfortunately, Linux is only beginning to support multiple visuals. As a result you have to make a choice when you configure your Linux workstation. You must decide if you want to be able to run pseudo-color applications, and continuously run out of your colormap, or to run 16,24, or 32-bit systems, which provide plenty of colors, but are unable to pseudo-color applications. Hopefully, in a year or two, Linux's multi-visual ability will be mature. For now I have added rudimentary support for 16, 24, and 32 bit depths. Because the overlay functionality (only available with pseudo-color) is fundamental to the operation of `Ricksep`, the hacks I introduced degrade performance.

CONCLUSIONS

Several new features have been added to the old `Rickmovie` program, the new `Ricksep`. Among these are expanded picking ability, annotation ability, multiple views, and support for multiple X11 visuals.

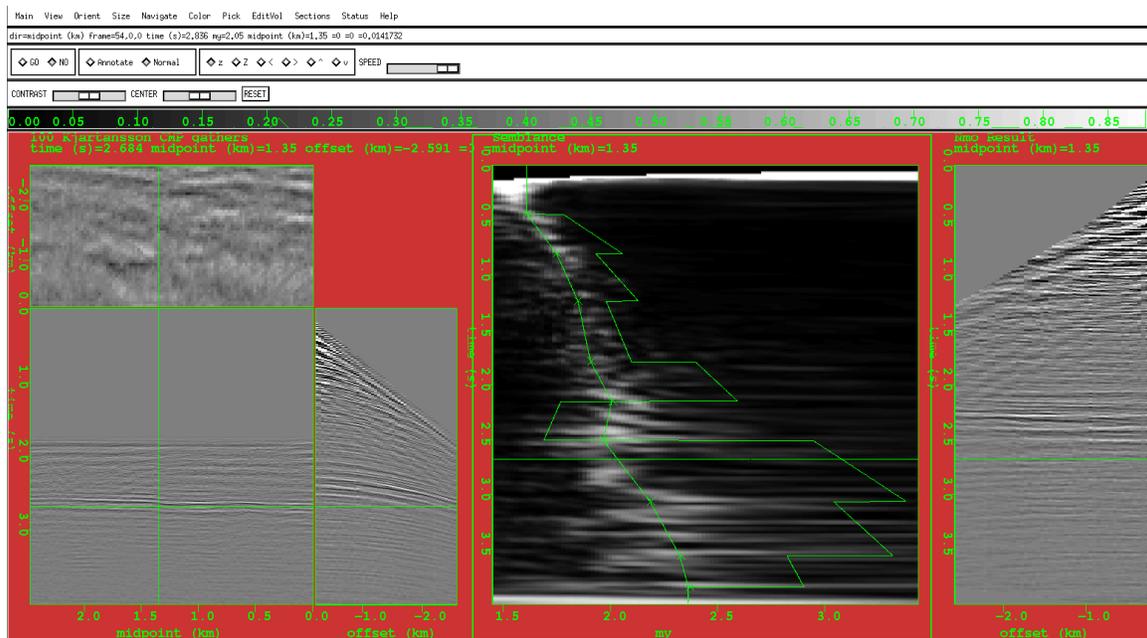


Figure 7: The Ricksep view when doing velocity analysis. The left panel is the input CMP sorted data. The center panel is the semblance, with the two curves representing RMS and interval velocity. The right panel shows the correct CMP gather. `bob2-rickvel` [NR]

REFERENCES

- Biondi, B., and van Trier, J., 1993, Visualization of multi-dimensional seismic datasets with CM-AVS: SEP-79, 1-12.
- Claerbout, J. F., 1981, On-line movies: SEP-28, 1-4.
- Claerbout, J. F., 1992, Earth Soundings Analysis: Processing versus Inversion: Blackwell Scientific Publications.
- Clapp, R. G., and Biondi, B., 1994, Iterative velocity model building for 3-D depth migration by integrating GOCAD and AVS: SEP-80, 635-644.
- Clapp, R. G., Biondi, B., and Karrenbach, M., 1994, AVS as a 3-D seismic data visualizing platform: SEP-82, 97-106.
- Clapp, R. G., 1995, SEP AVS User Guide: SEP-84, 395-408.
- Cole, S., and Dellinger, J., 1989, Vplot: SEP's plot language: SEP-60, 349-390.
- Dellinger, J., 1989, Why does SEP still use vplot: SEP-61, 327-336.
- Mora, C. B., Clapp, R. G., and Biondi, B., 1995, Velocity model building in avs: SEP-89, 133-144.

Mora, C. B., Clapp, R. G., and Biondi, B., 1996, Visualization of irregularly sampled seismic data with AVS: SEP-**93**, 75–86.

Ottolini, R., and Rocca, F., 1986, Movies of data lag histograms with application to deconvolution research: SEP-**50**, 35–42.

Ottolini, R., 1983, Movie cubes: SEP-**35**, 235–240.

Ottolini, R., 1988, Movies on the Macintosh II: SEP-**59**, 255–268.

Ottolini, R., 1990, Seismic movies on the XView graphics system: SEP-**65**, 315.

Sword, C. H., 1981, SEP goes to the movies: SEP-**28**, 11–20.

Short Note

Using the NTSC color space to double the quantity of information in an image

*Ioan Vlad*¹

INTRODUCTION

Geophysical images are, by their nature, intensity images: matrices of real numbers. Thus, representing an image in a color scale will only enlarge the visible dynamic color range without adding any information. But color can be used to encode information taken from a second geophysical image. The combination of the two images produces a meaningful image that can be understood better if two criteria are fulfilled: 1) The information used to create color (referred to from now on as the chrominance image) is entirely separated from the shading (black and white) information (referred to from now on as the luminance image). 2) The frequency content of the chrominance image must be different from that of the luminance image, preferably richer in lower frequencies, since the human eye is less sensitive to high spatial frequency color information. This short note will show how to combine two intensity images (i.e. a seismic section and a velocity model) into a single color image, using the NTSC (National Television Standards Committee) color space.

NTSC AND ITS PROPERTIES

The human eye has receptors for three primary colors - red, green and blue (Ruston, 1975). Thus, all color spaces have three dimensions. There are several color spaces available, the most widely known being RGB (Red-Green-Blue), HSV (Hue-Saturation-Value) and NTSC. NTSC is the color space used for television broadcast in the United States, and the only space among the above-mentioned ones that realizes a complete separation between the luminance and the chrominance information. NTSC has this property because when it was introduced, it had to separate the information used by the monochrome TV receivers from the supplementary one used by color receivers. (Blinn, 1993). The components of the NTSC color space are Y (the luminance component), I (the cyan-orange component), and Q (the green-purple component). There is a link between I and Q and the more intuitive hue and saturation: components I and Q have the property that, when modulated in quadrature in respect to each other by the 3.58

¹email: nick@sep.stanford.edu

MHz TV carrier wave and then summed, they form a signal (chrominance) whose magnitude represents the color saturation and whose phase represents the hue (Kuhn, 1996).

Saturation and hue are directly present as components of the HSV color space, and their meaning is more intuitive than that of the NTSC I and Q axes, so why not use HSV instead then? We must not not confuse between meaningfulness of the axes and meaningfulness of the object represented using these axes: because the separation between the luminance and chrominance information is not complete in HSV, those parts of the luminance image that are dark are not colored effectively and are rendered deep tones of gray, and ultimately a black pixel in the V component of HSV results in a black spot in a final image, thus not in the desired color. HSV images with the V component given by the luminance, and the H and S by the chrominance, will thus lack the ability to render fine detail, and the continuity of the color field will be broken by the dark pixels. The meaning of the dimensions of the HSV color space is easy to understand, but the final target, an image produced by taking the V component from one intensity image and the H and S components from the second intensity image, is harder to interpret than an image created in a similar manner using the NTSC color space. The exact way of constructing such an image using the NTSC color space is detailed in the next section.

CREATING THE COLOR IMAGE

The input data in this process consists of the luminance and the chrominance images. Both are stored in real-valued, normalized matrices (respectively L and C). The output will be a RGB image with three components: R,G and B, obtained by transforming a NTSC image composed of the matrices Y, I and Q. The first step is thus transforming L and C into the NTSC components Y, I and Q through a colormap operator (actually, a set of three related operators)

$$Y = M_1 L \quad (1)$$

$$I = M_2 C \quad (2)$$

$$Q = M_3 C \quad (3)$$

The next section will elaborate on the construction of the three operators. Once the Y, I and Q matrices are obtained, we basically have the color image, but since computer devices work with RGB, we usually need to transform to RGB with the equation

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.000 & 0.956 & 0.621 \\ 1.000 & -0.272 & -0.647 \\ 1.000 & -1.106 & 1.703 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix} \quad (4)$$

before writing the file to disk (Kuhn, 1996).

BUILDING THE THREE COLORMAP OPERATORS

In simple cases M_1 , from eq. (1), is a histogram equalization operator (for contrast adjustment) or only the identity matrix, and M_2 and M_3 represent raising each point in the matrix to a

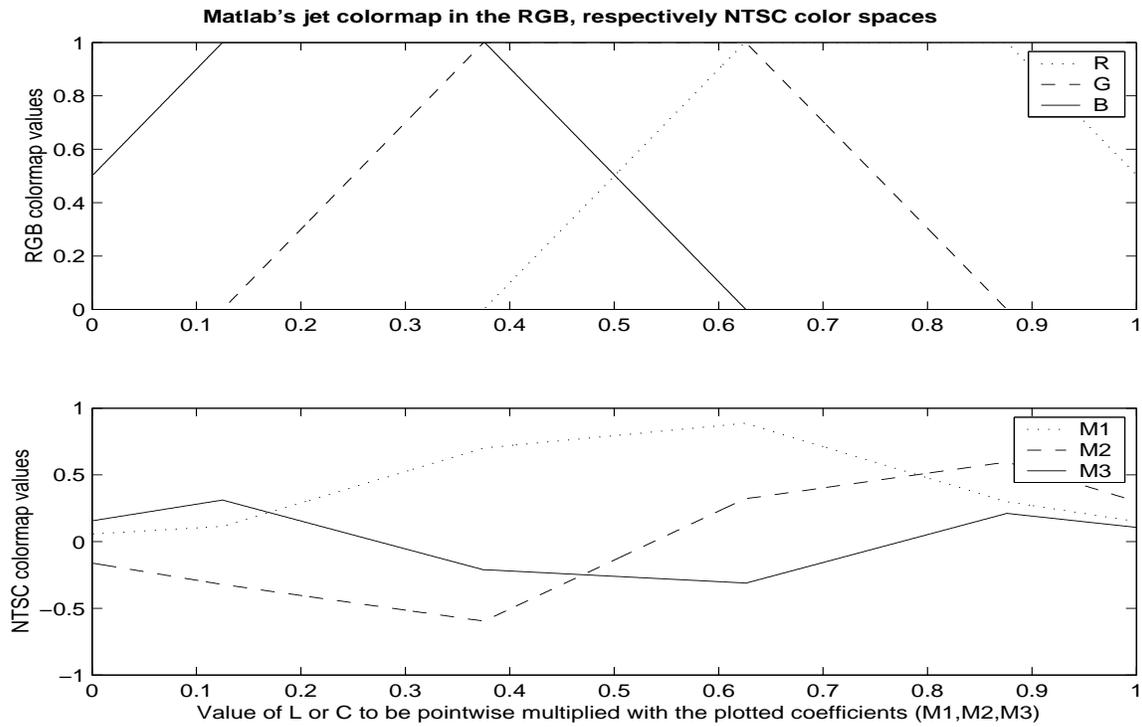


Figure 1: The main part of the operators M_1 , M_2 and M_3 consists of the transformation from RGB to NTSC of a common colormap `nick1-jetcolormap` [CR]

power. Several variations that would result in monochrome colormaps suitable especially for highlighting linear, coherent shapes (fig. 6) can be imagined. If a more complex colormap is desired, the three operators must be constructed by transforming common RGB colormaps to NTSC, with the equation:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

An example is provided in Figure 1 for Matlab's jet colormap.

APPLICATIONS

The use of color for highlighting linear objects is illustrated by figure 6. (All the color images were grouped at the end of the paper due to printing necessities). Luminance is given by a simple synthetic reflectivity model, and the result of applying a local monoplane annihilator to the luminance image constitutes the chrominance (fig 2). The two images were generated with code from Claerbout (2001). It must be noticed that the colored pixels do not obscure the local features of the image (the highlighting is not "drawn over" the luminance image), but they have the shading given by the synthetic model itself.

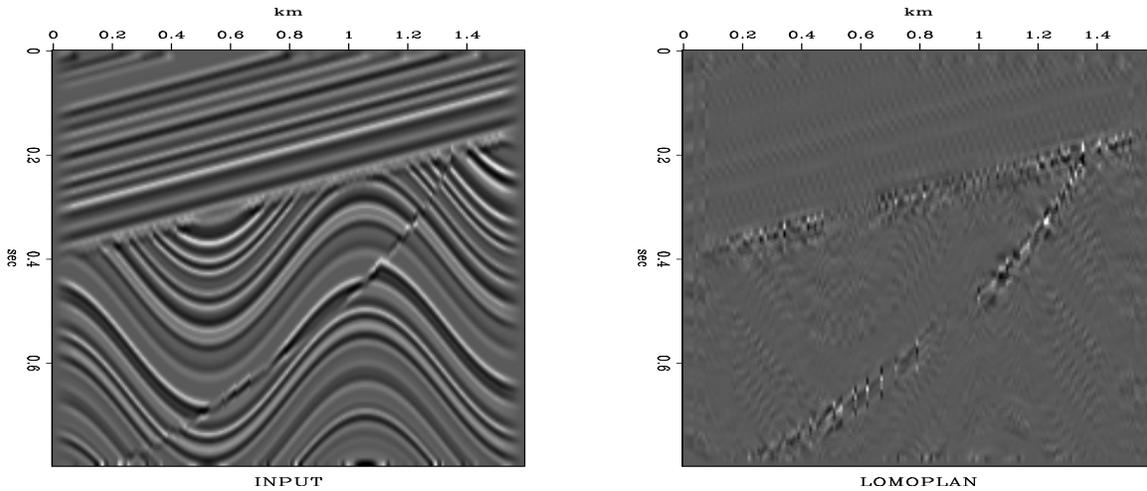


Figure 2: Luminance (left) and chrominance (right) for Fig. 6 `nick1-lomoplan` [ER]

One of the best uses of the tool described in this paper is representing both the velocity model and the seismic section in the image. Fig. 3 represents the velocity model (chrominance), Fig. 4 the seismic section, and Fig. 7 the result of applying the above described algorithm to the two images. Operator M consisted in this case of a histogram equalization for the luminance, then of a transformation to Y, I and Q through the components of the jet colormap (fig. 1).

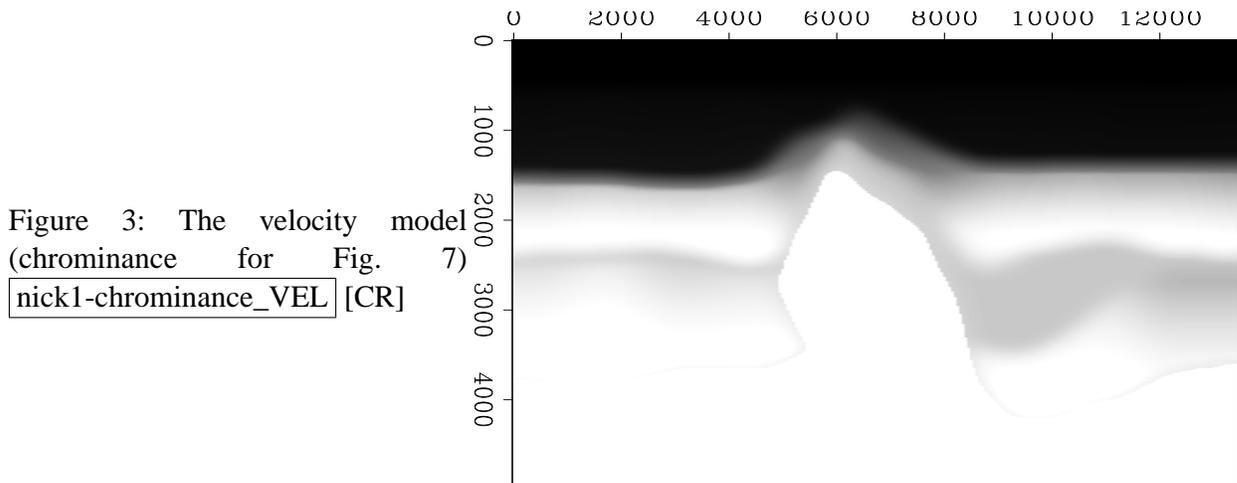


Figure 3: The velocity model (chrominance for Fig. 7) `nick1-chrominance_VEL` [CR]

Dip fields, described in Fomel (2000) can be helpful in interpreting seismic data. Fig. 5 shows the value of the strongest dip in the corresponding point of Fig. 4, computed with the programs accompanying Fomel (2000). It is not easy to visually correlate the features in the two images. But when the dip field is used for chrominance, and the seismic image for luminance (with the operator M containing a 97 percent clip before applying jet colormap), the unconformities, as well as the areas with conflicting dips, become highly visible. (Fig. 8)

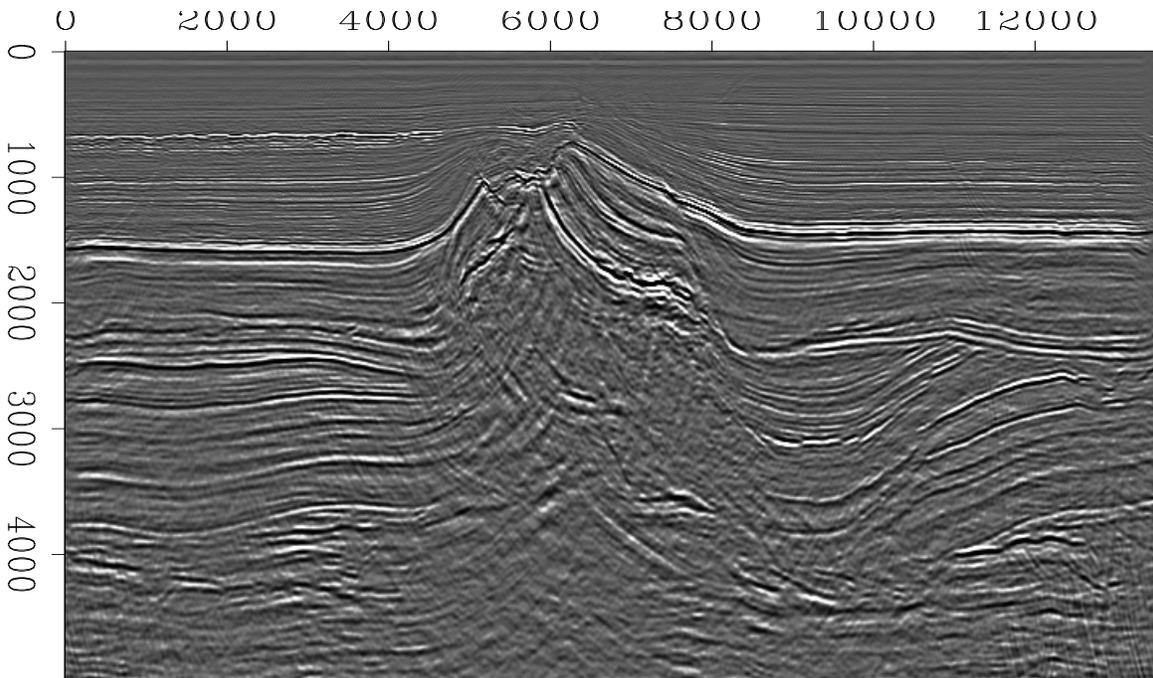


Figure 4: Luminance for Fig. 7 `nick1-luminance_VEL` [CR]

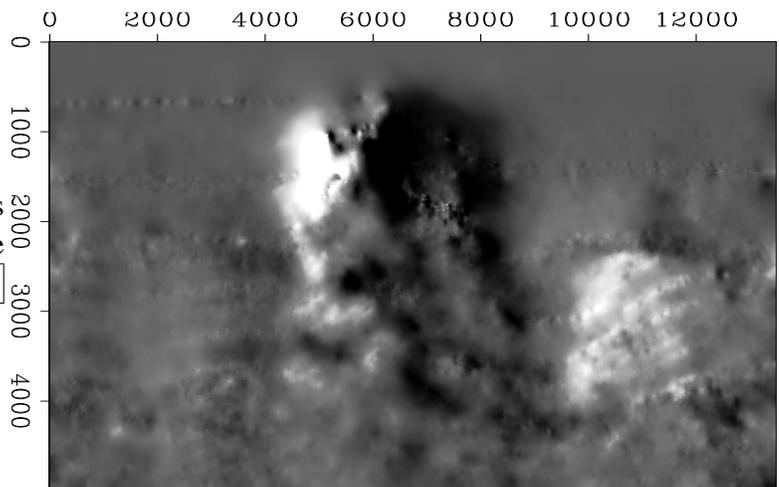


Figure 5: Dip field of Fig. 4. This image will be used as chrominance for Fig. 8 `nick1-chrominance_DIP` [CR]

CONCLUSIONS

NTSC is the color space with the best separation between the luminance and the chrominance information. It is useful for enhancing the interpretability of geophysical images in a simple, intuitive way, easy to implement into software and computationally inexpensive. Several other examples can be found at <http://sepwww.stanford.edu/sep/nick/research/NTSC/> . Future work include implementation of this feature into SEPlib.

REFERENCES

Blinn, J. F., 1993, Ntsc: Nice technology, super color: IEEE Computer Graphics and Applications, **13**, no. 2, 17–23.

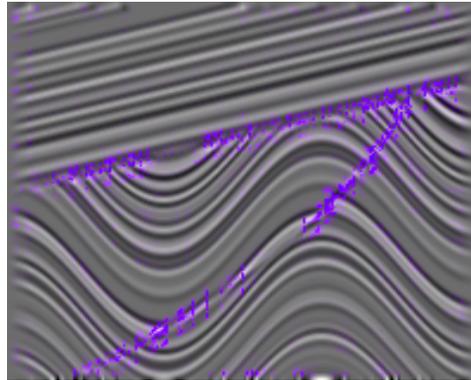
Claerbout, J. Geophysical estimation by example: http://sepwww.stanford.edu/sep/prof/gee/toc_html/index.html, 2001.

Fomel, S., 2000, Applications of plane-wave destructor filters: SEP-105, pages 1–27.

Kuhn, K. J., 1996, Conventional analog television - an introduction: <http://www.ee.washington.edu/conselec/CE/kuhn/ntsc/95x4.htm>.

Ruston, W., 1975, Visual pigments and color blindness: Scientific American, pages 64–74.

Figure 6: Image built using NTSC, the color highlights the discontinuity in the seismic section
`nick1-lomocolor` [CR]



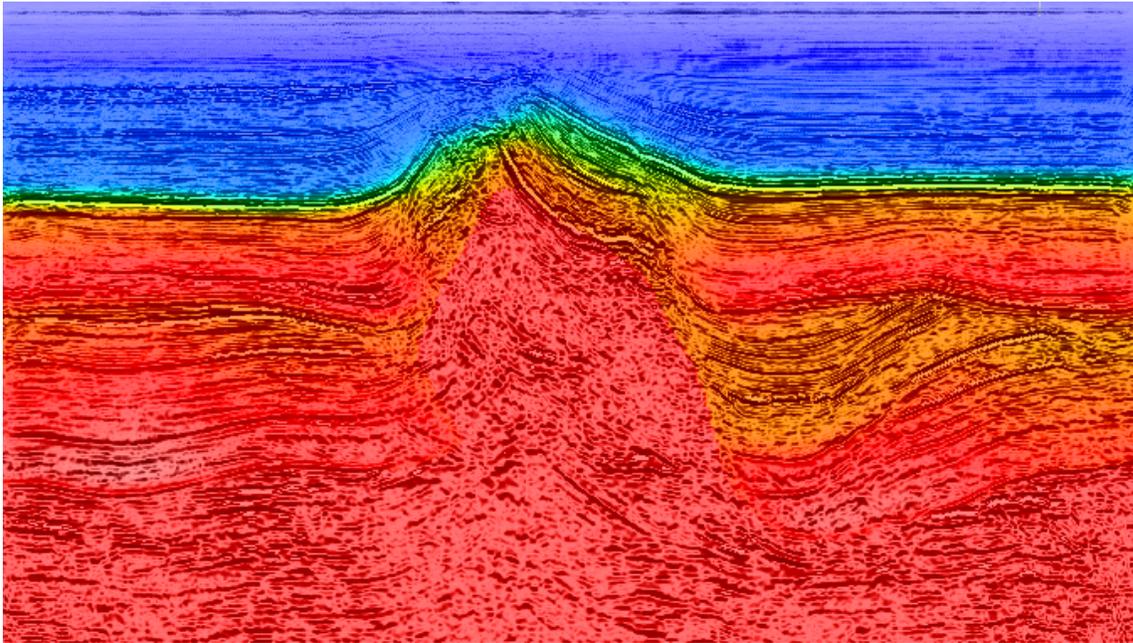


Figure 7: Composite image with shading taken from fig. 4 and color from Fig. 3. Electronic viewing of the image is highly recommended in order to show the full power of the technique described in this paper. `nick1-color_VEL` [CR]

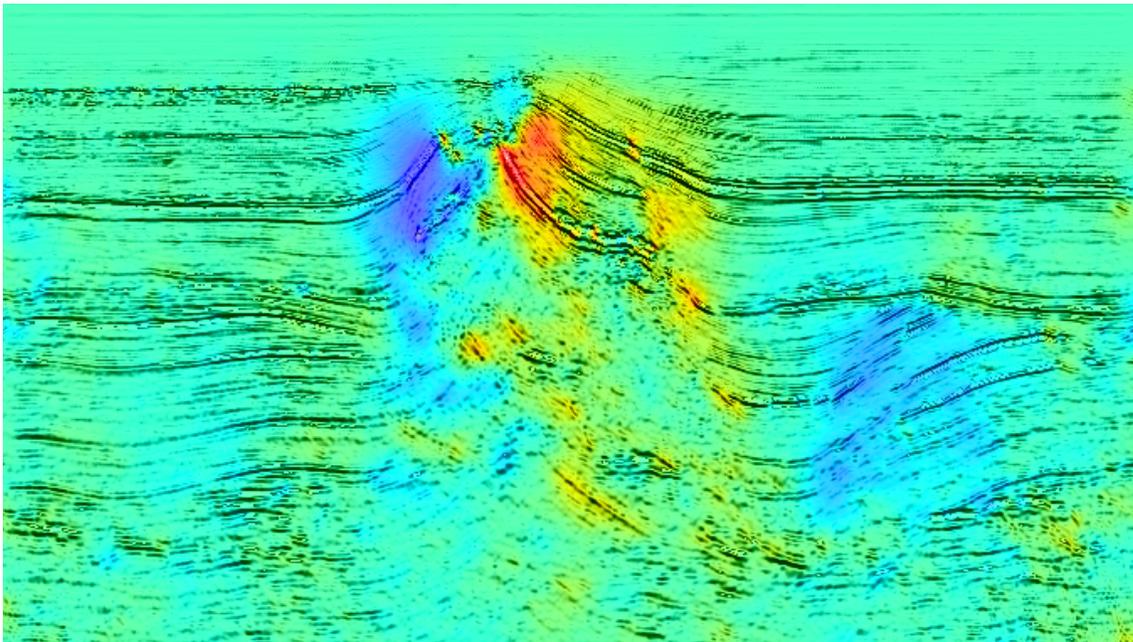


Figure 8: Coloring the dips on the seismic section: composite image with shading taken from Fig. 4 and color from fig. 5. Electronic viewing of the image is highly recommended in order to show the full power of the technique described in this paper. `nick1-color_DIP` [CR]

SEP manual

Robert Clapp, Marie Prucha, Paul Sava, Joe Dellinger, and Biondo Biondi¹

ABSTRACT

INTRODUCTION

CONCLUSIONS

REFERENCES

¹**email:** paul@sep.stanford.edu

SEP ARTICLES PUBLISHED OR IN PRESS

- Alkhalifah, T., and Fomel, S., 2001, Implementing the fast marching eikonal solver: Spherical versus Cartesian coordinates: *Geophysical Prospecting*, accepted for publication.
- Alkhalifah, T., Fomel, S., and Biondi, B., 2001, The space-time domain: Theory and modelling for anisotropic media: *Geophysical Journal International*, **144**, 105–113.
- Biondi, B., 2000, Stable wide-angle Fourier-finite difference downward extrapolation of 3-D wavefields: submitted for publication in *Geophysics*.
- Biondi, B., 2001a, Kirchhoff imaging beyond aliasing: *Geophysics*, in press.
- Biondi, B., 2001b, Stable wide-angle Fourier-finite difference downward extrapolation of 3-D wavefields: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Chemingui, N., and Biondi, B., 2000, Seismic data reconstruction by inversion to common offset: submitted for publication in *Geophysics*.
- Clapp, R., 2001, Geologically constrained migration velocity analysis: field example: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Fomel, S., and Grechka, V., 2001, Nonhyperbolic reflection moveout of p -waves: An overview and comparison of reasons: submitted for publication in *Geophysical Prospecting*.
- Fomel, S., and Vaillant, L., 2001, Evaluating the Stolt-stretch parameter: *Journal of Seismic Exploration*, accepted for publication.
- Fomel, S., Clapp, R., Prucha, M., and Berryman, J., 2001, Iterative resolution estimation in least-squares Kirchhoff migration: submitted for publication in *Geophysical Prospecting*.
- Fomel, S., 2001a, Applications of plane-wave destructor filters: submitted for publication in *Geophysics*.
- Fomel, S., 2001b, Asymptotic pseudo-unitary stacking operators: submitted for publication in *Geophysics*.
- Fomel, S., 2001c, Differential offset continuation in theory: submitted for publication in *Geophysics*.
- Fomel, S., 2001d, Seismic reflection data interpolation with differential offset and shot continuation: submitted for publication in *Geophysics*.
- Guitton, A., Brown, M., Rickett, J., and Clapp, R., 2001, Multiple attenuation using a t-x pattern-based subtraction method: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Guitton, A., 2001, Coherent noise attenuation using inverse theory and prediction error filters: 63rd Mtg., Eur. Assoc. Geosc. Eng., Abstracts, accepted for presentation.

- Rickett, J. E., and Claerbout, J. F., 2000, Calculation of the acoustic solar impulse response by multi-dimensional spectral factorization: *Solar Physics*, **92**, no. 1/2, 203–210.
- Rickett, J. E., and Lumley, D. E., 2001, Cross-equalization processing for time-lapse seismic reservoir monitoring data: A case study from the Gulf of Mexico: *Geophysics*, accepted for publication.
- 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, submitted.
- Rickett, J., Guitton, A., and Gratwick, D., 2001, Adaptive multiple subtraction with non-stationary helical shaping filters: 63rd Mtg., Eur. Assoc. Geosc. Eng., Abstracts, accepted for presentation.
- Rickett, J., 2001, Model-space vs. data-space normalization for finite-frequency depth migration: 71th Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Rosales, D., and Rickett, J., 2001, *PS*-wave polarity reversal in angle domain common-image gathers: 71st Ann. Internat. Meeting, Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Sava, P., and Fomel, S., 2001a, 3-D travelttime computation by Huygens wavefront tracing: *Geophysics*, **66**, no. 3, in press.
- Sava, P., and Fomel, S., 2001b, Migration angle-gathers by Fourier transform: submitted for publication in *Geophysical Prospecting*.
- Sava, P., Biondi, B., and Fomel, S., 2001, Amplitude-preserved common image gathers by wave-equation migration: 71st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, submitted.
- Sava, P., 2001, Prestack residual migration in the frequency domain: submitted for publication in *Geophysics*.
- Schwab, M., Karrenbach, M., and Claerbout, J., 2000, Making scientific computations reproducible: *Computing in Science and Engineering*, **2**, no. 6, 61–67.

SEP PHONE DIRECTORY

Name	Phone	Login Name
Alvarez, Gabriel	723-0463	gabriel
Artman, Brad	723-6007	brad
Berryman, James	723-1250	berryman
Biondi, Biondo	723-1319	biondo
Brown, Morgan	723-6006	morgan
Claerbout, Jon	723-3717	jon
Clapp, Robert	723-0253	bob
Curry, William	725-1334	bill
Guitton, Antoine	723-6007	antoine
Karpushin, Andrey	725-1334	andrey
Lau, Diane	723-1703	diane
Prucha, Marie	723-0253	marie
Rosales, Daniel	723-6006	daniel
Sava, Paul	723-0463	paul
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 FOR 2000-2001

Name	Company	Telephone	E-Mail
Raymond Abma, Co-Chair	BP Amoco	(281) 366-4604	abmar1@bp.com
Francois Audebert, Co-Chair	CGG	33 (1) 64473324	faudebert@cgg.com
Biondo Biondi	SEP	(650) 723-1319	biondo@sep.stanford.edu
Luis Canales	WesternGeco	(713) 689-5717	luis.canales@westgeo.com
Jon Claerbout	SEP	(650) 723-3717	jon@sep.stanford.edu
Stewart Levin	Landmark Graphics	(303) 265-8649	salevin@lgc.com
Reinaldo Michelena	PDVSA Intevp	58 (212) 9086855	michelena@pdvsa.com
Simon Spitz	CGG	(281) 646-2426	sspitz@cgg.com

Research Personnel

Brad Artman received his B.Sc. in Geophysical Engineering from the Colorado School of Mines in December 1996. He worked at Shell Deepwater Development Company in New Orleans in petrophysical and geophysical capacities until joining SEP in the fall of 2000 to work toward a Ph.D. Brad is member of SEG and SPWLA.



James G. Berryman received a B.S. degree in physics from Kansas University (Lawrence) in 1969 and a Ph.D. degree in physics from the University of Wisconsin (Madison) in 1975. He subsequently worked on seismic prospecting at Conoco. His later research concentrated on sound waves in rocks – at AT&T Bell Laboratories (1978-81) and at Lawrence Livermore National Laboratory (1981-), where he is currently a physicist in the Energy and Environment Directorate. Continuing research interests include seismic and electrical methods of geophysical imaging and waves in porous media containing fluids. He is a member of ASA, AGU, APS, and SEG.



Biondo L. Biondi graduated from Politecnico di Milano in 1984 and received an M.S. (1988) and a Ph.D. (1990) in geophysics from Stanford. SEG Outstanding Paper award 1994. During 1987 he worked as a Research Geophysicist for TOTAL, Compagnie Francaise des Petroles in Paris. After his Ph.D. at Stanford Biondo worked for three years with Thinking Machines Co. on the applications of massively parallel computers to seismic processing. After leaving Thinking Machines Biondo started 3DGeo Development, a software and service company devoted to high-end seismic imaging. Biondo is now Associate Professor (Research) of Geophysics and leads SEP efforts in 3-D imaging. He is a member of SEG and EAGE.



Morgan Brown received a B.A. in Computational and Applied Mathematics from Rice University in 1997 and is currently completing a Ph.D. at SEP. Morgan worked as a research intern with Western Geophysical in 1997 and with Landmark Graphics in 2000. He is a member of SEG.



Robert Clapp received his B.Sc.(Hons.) in Geophysical Engineering from Colorado School of Mines in May 1993. He joined SEP in September 1993, received his Masters in June 1995, and his Ph.D. in December 2000. He is a member of the SEG and AGU.



Barbara E. Cox received her M.Sc. in Geophysics from Utrecht University in November 1998. She joined the DELPHI consortium (Delft University of Technology, Department of Applied Physics) in June 1999, where she is working towards a Ph.D. in physics. She is a member of the SEG and the EAGE. She has been visiting SEP during the summer of 2001.



William Curry is a second year student with SEP, who graduated in 2000 from the University of Alberta with a B.Sc. (Hons.) in Geophysics. From 1998-1999, he worked as a seismic interpreter with Jet Energy Corp. in Calgary, Canada. Bill is a student member of the AGU, CGU, CSEG, EAGE and SEG.



Andrey Karpushin received his Diploma in Geophysics from the Novosibirsk State University, Russia in 1993. After the graduation he worked at the Novosibirsk Institute of Geology and Geophysics. Before coming to SEP he worked at Dowell Schlumberger. He joined SEP in 2000 and now he is currently working toward a Ph.D. in geophysics. He is a member of SEG.



Clement Kostov graduated from Ecole des Mines de Paris in 1983, then received a MSc (1985) in geostatistics and a PhD in geophysics, both from Stanford. Since 1990, he has been working with Schlumberger, Geco-Prakla, and WesternGeco in research, development and technical marketing positions in France, UK and USA. Currently Clement is with the R&D Imaging Research group of WesternGeco in Houston. In summer 2001, Clement visited SEP for a month.



Marie Prucha received her B.Sc.(Hons.) in Geophysical Engineering from Colorado School of Mines in May 1997. She joined SEP in September 1997 and received her MS in June 1999. She is currently working towards a Ph.D. in geophysics. She is a member of SEG.



Paul Sava graduated in June 1995 from the University of Bucharest, with an Engineering Degree in Geophysics. Between 1995 and 1997 he was employed by Schlumberger GeoQuest. He joined SEP in 1997, received his M.Sc. in 1998, and continues his work toward a Ph.D. in Geophysics. His main research interest is in seismic imaging using wave-equation techniques. He is a member of SEG, EAGE and the Romanian Society of Geophysics.



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 currently working towards a Ph.D. in geophysics at Stanford. He is a member of SEG.



SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2000-2001

Amerada Hess Corporation
 One Allen Center
 500 Dallas St.
 Houston, TX 77002
 U.S.A.
 tel: (713) 609-5829
 fax: (713) 609-5999
 contact: Scott Morton
 email: morton@hess.com

BP Amoco Corporation
 200 West Lake Park Blvd.
 1018 WL4
 Houston, TX 77079
 U.S.A.
 tel: (281) 366-3611
 fax: (281) 366-5856
 contact: John Etgen
 email: jetgen@amoco.com

CGG Americas, Inc.
 16430 Park Ten Place
 Houston, TX 77084
 U.S.A.
 tel: (281) 646-2502
 fax: (281) 646-2622
 contact: Simon Spitz
 email: sspitz@us.cgg.com

Chevron Petroleum Technology Co.
 6001 Bollinger Canyon Rd., Bldg. D
 P.O. Box 6019
 San Ramon, CA 94583-0719
 U.S.A.
 tel: (925) 842-6232
 fax: (925) 842-2076
 contact: Ray Ergas
 email: rae@chevron.com

Chinese Petroleum Corporation
 Geophysical Data Processing Ctr., OPED
 8, Lane 22, Tunghwa St., Section 1
 Peitou District
 Taipei, Taiwan 11206
 Republic of China
 tel: 886 (2) 2821 6313
 fax: 886 (2) 2821 3147
 contact: Jenyang Lin
 email: jenyang@gdpc.cpcoped.com.tw

Conoco Inc.
 Seismic Imaging Technology
 P.O. Box 1267
 1000 S. Pine
 Ponca City, OK 74602-1267
 U.S.A.
 tel: (580) 767-2046
 fax: (580) 767-2887
 contact: Alan R. Huffman
 email: alan.r.huffman@usa.conoco.com

Ecopetrol-ICP
 Laboratorio de Geofisica
 A.A. 4185 Bucaramanga
 Colombia
 tel: (57) 76 445420
 fax: (57) 76 445444
 contact: Alfredo Tada
 email: atada@ecopetrol.com.co

ENI SPA - Agip Division
 Dept. RIGE
 via Unione Europea 3
 I-20097 S. Donato Milanese
 Italy
 tel: 39 (02) 520 55308
 fax: 39 (02) 520 35132
 contact: Vittorio De Tomasi
 email: vittorio.detomasi@agip.it

SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2000-2001

ExxonMobil Upstream Research Co.
 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 Dickens
 email: tadicke@upstream.xomcorp.com

4th Wave Imaging Corporation
 Suite 200, 16A Journey
 Aliso Viejo, CA 92656
 U.S.A.
 tel: (949) 916-9787
 fax: (949) 916-9786
 contact: David Lumley
 email: david.lumley@4thwaveimaging.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: Takeshi Ikawa
 email: ikawa@jgi.co.jp

Landmark Graphics Corporation
 7409 S. Alton Court
 Suite 100
 Englewood, CO 80112-2301
 U.S.A.
 tel: (303) 779-8080
 fax: (303) 796-0807
 contact: Burke G. Angstman
 email: bangstman@lgc.co

Linux Networx Inc.
 8689 South 700 West
 Sandy, UT 84070
 U.S.A.
 tel: (801) 304-0226
 fax: (801) 568-1010
 contact: William Harman
 email: wharman@lnxi.com

Norsk Hydro ASA
 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

PDVSA Intevep
 Urb. Santa Rosa, Sector El Tambor
 Los Teques, Miranda
 Venezuela
 tel: 58 (212) 908 6855
 fax: 58 (212) 908 7078
 contact: Reinaldo Michelena
 email: michelenar@pdvsa.com

SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2000-2001

Petrobras
 Av. Chile 65, Depex/Ditrex
 Centro, Rio de Janeiro
 20132 RJ
 Brazil
 tel: 55 (21) 534 2706
 fax: 55 (21) 534 1076
 contact: Carlos A. Cunha Filho
 email: s002@ep.petrobras.com.br

PGS, Inc.
 738 Highway 6 South
 Suite 300
 Houston, TX 77079
 U.S.A.
 tel: (281) 589-6725
 fax: (281) 589-6558
 Contact: James R. Myron
 email: jim@hstn.seres.pgs.com

Phillips Petroleum Company
 500A Plaza Office Bldg.
 Bartlesville, OK 74004
 U.S.A.
 tel: (918) 661-9425
 fax: (918) 661-5250
 contact: Dan Whitmore
 email: ndwhitm@ppco.com

Saudi Aramco Oil Company
 Geophysical R&D
 EXPEC BLDG, X-3480A
 Dhahran 31311
 Saudi Arabia
 tel: 966 (3) 874 7262
 fax: 966 (3) 873 1020
 contact: Mohammed N. Alfaraj
 email: farajmn@mail.aramco.com.sa

Schlumberger Cambridge Research
 High Cross
 Madingley Road
 Cambridge, England CB3 0EL
 United Kingdom
 tel: 44 (1223) 325 337
 fax: 44 (1223) 361 473
 contact: Dave Nichols
 email: nichols@cambridge.scr.slb.com

Shell E&P Technology Company
 Bellaire Technology Center
 3737 Bellaire Blvd.
 Houston, TX 77025
 U.S.A.
 tel: (713) 245-7285
 fax: (713) 245-7339
 contact: Chris Corcoran
 email: ctcorcoran@shellus.com

Texaco Group Inc.
 P.O. Box 5060
 5901 S. Rice Ave.
 Bellaire, TX 77402-5060
 U.S.A.
 tel: (713) 432-6798
 fax: (713) 432-6908
 contact: Youngsun Shin
 email: shiny@texaco.com

3DGeo Development Inc.
 465 Fairchild Drive
 Suite 226
 Mountain View, CA 94043-2251
 U.S.A.
 tel: (650) 969-3886
 fax: (650) 969-6422
 contact: Dimitri Bevc
 email: dimitri@3dgeo.com

SPONSORS OF THE STANFORD EXPLORATION PROJECT, 2000-2001

TotalFinaElf
Bureau M/059, CSTJF
Avenue Larribau
64018 Pau Cedex
France
tel: 33 (559) 836 786
fax: 33 (559) 834 858
contact: M. Henri Calandra
email: henri.calandra@totalfinaelf.com

Veritas DGC Ltd.
Crompton Way
Manor Royal Estate
Crawley, West Sussex RH10 2QR
United Kingdom
tel: 44 (1293) 443219
fax: 44 (1293) 443012
contact: Helmut Jakubowicz
email: helmut_jakubowicz@veritasdgc.com

Unocal (Spirit Energy 76)
E&P Technology
14141 Southwest Freeway
Sugar Land, TX 77478
U.S.A.
tel: (281) 287-7481
fax: (281) 287-5110
contact: Phil Schultz
email: phil.schultz@unocal.com

WesternGeco
10001 Richmond Ave.
Houston, TX 77242-4299
U.S.A.
tel: (713) 689-5717
fax: (713) 689-5757
contact: Luis Canales
email: luis.canales@westgeo.com