Applying Compactness Constraints To Differential Traveltime Tomography

J.B. Ajo-Franklin,
Earth Resources Laboratory
Dept. of Earth, Atmospheric, and Planetary Sciences
Massachusetts Institute of Technology
Cambridge, MA 02142

B.J. Minsley,
Earth Resources Laboratory

T.M. Daley,
Earth Science Division,
Lawrence Berkeley National Laboratory, CA, USA

May 7, 2007

Abstract

Tomographic imaging problems are typically ill-posed and often require the use of regularization techniques to guarantee a stable solution. Minimization of a weighted norm of model length is one commonly used secondary constraint. Tikhonov methods exploit low-order differential operators to select for solutions that are small, flat, or smooth in one or more dimensions. This class of regularizing functionals may not always be appropriate, particularly in cases where the anomaly being imaged is generated by a non-smooth spatial process. Timelapse imaging of flow-induced velocity anomalies is one such case; flow features are often characterized by spatial compactness or connectivity. By performing inversions on differenced arrival time data, the properties of the timelapse feature can be directly constrained. We develop a differential traveltime tomography algorithm which selects for compact solutions i.e. models with a minimum area of support, through application of model-space iteratively reweighted least squares. Our technique is an adaptation of minimum support regularization methods previously explored within the potential theory community. We compare our inversion algorithm to the results obtained by traditional Tikhonov regularization for two simple synthetic models; one including several sharp localized anomalies and a second with smoother features. We use a more complicated synthetic test case based on multiphase flow results to illustrate the efficacy of compactness constraints for contaminant infiltration imaging. We conclude by applying the algorithm to a CO_2 sequestration monitoring dataset acquired at the Frio pilot site. We observe that in cases where the assumption of a localized anomaly is correct, the addition of compactness constraints improves image quality by reducing tomographic artifacts and spatial smearing of target features.

[In Press : *Geophysics*, **2007]**

1 Introduction

The inversion of geophysical data, and tomographic imaging problems in particular, are often both non-unique and ill-posed. When we are confronted with a multitude of valid answers, all sensitive to small variations in noise, secondary constraints can be added to both stabilize the inversion and to select solutions which fulfill an independent notion of what a "good" model should look like. Regularization techniques accomplish both goals by minimizing a weighted semi-norm of solution length in addition to fitting the data. One such approach, originally developed by *Tikhonov and Arsenin* [1977], minimizes one or more low order (0th, 1st or 2nd) spatial derivatives of the model to help choose small, flat or smooth solutions. Despite the fact that neither flatness nor smoothness are intrinsic properties of the earth, Tikhonov methods have enjoyed remarkable success and are routinely applied to a wide range of parameter estimation problems [*Aster et al.*, 2005].

As discussed by *Constable et al.* [1987], the purpose of regularization in the inverse problem is to introduce stability while recovering models that do not contain more complicated features than can be justified by the data. Smoothness can be a desirable quality because it suppresses unnecessary model complexity, and often provides a reasonable representation of earth structures. Additionally, the application of a linear regularization operator on the model using an l_2 norm results in a quadratic term in the objective function that is minimized by solving a linear system of equations.

We advocate the selection of regularization operators which incorporate some notion of the physics responsible for observed property variations *e.g.* subsurface flow, thermal diffusion, or fracture propagation. Due to the complexity of these processes, heuristic constraints which select for models with related characteristics might be appropriate. Since flow processes tend to localize in zones of high permeability, regularization operators favoring compact or connected anomalies seem reasonable. The infiltration of dense non-aqueous contaminants [*Kueper et al.*, 1993] and the transport of saline tracers through permeable fractures [*Day-Lewis et al.*, 2003] are two examples of compact flow features suitable for geophysical characterization.

The dynamic nature of flow processes makes them amenable to timelapse methods which seek to delineate temporal changes in subsurface properties. Unfortunately, standard timelapse imaging techniques are incapable of directly constraining model perturbations; the most common approach is to perform independent inversion for datasets acquired at multiple times followed by model domain subtraction to yield a differenced image or a set of differenced attributes [*Greaves and Fulp*, 1987]. In this case, constraints can only be applied to each independent dataset but not to the difference between them. This distinction is crucial since the spatial characteristics of the dynamic process (*e.g.* flow) may have different geometric properties than background structures. An alternative approach is to subtract datasets acquired at multiple times directly in the data domain followed by inversion of the differenced data to recover model perturbations. This approach, which we refer to as differential or difference inversion, allows spatial constraints to be directly applied to changes in model properties.

Various regularization schemes have been introduced that provide inversion stability using constraints that promote simple, though not necessarily smooth, features in the model without introducing unnecessary complexity. Some examples of alternate regularization methods include minimization of the low order spatial derivatives of the model using an l_1 norm [*Claerbout and Muir*, 1979], minimizing the area occupied by model parameters [*Last and Kubik*, 1983] or their spatial derivatives [*Portniaguine and Zhdanov*, 1999], and minimization of the moment of inertia of an object [*Guillen and Menichetti*, 1984]. In all of these cases, the objective function is no longer quadratic, and solution of the non-linear inverse problem requires the use of model-space iteratively reweighted least squares (IRLS). IRLS techniques are more commonly used in the data domain to solve inverse problems in the l_p norm for $1 \leq p \leq 2$ [*Scales et al.*, 1988; *Bube and Langan*, 1997] but can be easily adapted to model-space reweighting [*Farquharson and Oldenburg*, 1998].

One particular application of model regularization that minimizes the l_1 norm of the model gradient is often referred to as total variations. This approach has been used for image processing where reconstruction of sharp edges is required [*Rudin et al.*, 1992; *Acar and Vogel*, 1994], as well as geophysical inverse problems where the target of interest is not inherently smooth [*Bertete-Aguirre et al.*, 2002; *Yu and Dougherty*, 2000]. The l_1 measure of the model gradients does not penalize sharp boundaries as strongly as the l_2 measure, therefore allowing for models that are more blocky.

Compact body inversion, as developed by *Last and Kubik* [1983], is one approach used in the potential field community for selecting compact models while still satisfying data misfit constraints. In this case, compactness implies a solution which minimizes the area of an anomaly in 2D or the volume of an anomaly in 3D. This concept is further developed by *Portniaguine and Zhdanov* [1999] to select models where the spatial gradients of an anomaly, rather than the anomaly itself, are compact. They use the term minimum gradient support (MGS) to describe this regularization method. *Youzwishen and Sacchi* [2006] use this approach to reconstruct blocky acoustic velocity models from synthetic seismic datasets. Both compactness and MGS allow for sharper model variations than traditional smoothness constraints, though MGS tends to produce more blocky images due to the penalty on model gradients.

Given a variety of regularization techniques that provide inversion stability without introducing unnecessary solution complexity, one should choose a method that is consistent with the expected physical properties of the model. This choice is equivalent to introducing prior knowledge into the inverse problem. Timelapse seismic traveltime tomography, a technique with demonstrated utility in a monitoring context (e.g. Lazaratos and Marion, 1997), is one of many geophysical inverse problems which might benefit from the incorporation of compactness constraints. As mentioned previously, geophysical perturbations induced by flow processes often yield features with spatially localized properties suggesting that compactness is an appropriate metric with which to evaluate solutions. In this document, we will pose the differential traveltime tomography problem in the formalism of *Last and Kubik* [1983] and demonstrate the resulting algorithm on a simple synthetic test problem, a more realistic problem based on contaminant imaging, and on a crosswell seismic monitoring dataset acquired at the Frio pilot sequestration site.

1.1 Principles & Theory

We initially consider the general linear inverse problem where a linear operator, $\mathbf{G}$, maps a model ($\mathbf{m}$) to a dataset ($\mathbf{d}$) i.e. $\mathbf{G}\,\mathbf{m} = \mathbf{d}$. Traditional Tikhonov regularization selects solutions that minimize an objective function, $\Phi(\mathbf{m})$, combining a measure of data misfit and a weighted semi-norm of model length in the l_2 sense,

$$\Phi(\mathbf{m}) = \|\mathbf{G}\,\mathbf{m} - \mathbf{d}\|_2^2 + \lambda^2 \|\mathbf{W}\,\mathbf{m}\|_2^2, \tag{1}$$

where $\mathbf{W}$ is typically either $\mathbf{I}$ or a low order differential operator and λ, referred to as the regularization parameter, allows the weight given to solution length to vary. When $\mathbf{W}$ is a 1st spatial derivative operator, bias is given towards flat models while use of a laplacian or split 2nd derivative operators favors smooth models. Minimizing equation 1 results in an augmented least-squares problem of the form,

$$\begin{bmatrix} \mathbf{G} \\ \lambda \mathbf{W} \end{bmatrix} \mathbf{m} = \begin{bmatrix} \mathbf{d} \\ 0 \end{bmatrix}. \tag{2}$$

While Tikhonov schemes have enjoyed successful application in many fields, as mentioned previously they rely on a somewhat arbitrary choice of prior structure to stabilize the inversion problem. *Last and Kubik* [1983] developed an alternative regularization strategy which selects for models with causative bodies of minimum area in addition to fitting the data. A key aspect of such a strategy is a consistent definition of area in the context of imaged anomalies. *Last and Kubik* [1983] introduced an area metric, $A(\mathbf{m})$, for n elements of constant size which can be written as,

$$A(\mathbf{m}) = a_e \lim_{\beta \to 0} \sum_{i=1}^{n} \frac{m_i^2}{m_i^2 + \beta}, \tag{3}$$

where a_e is the area of a single element, m_i is the ith model parameter, and β is a factor to remove the singularity in cases where $m_i \to 0$. In the limit of small β, the interior of the right hand side of equation 3 evaluates to 0 for cases where $m_i = 0$ and 1 for non-zero values. Equation 3 can thus be viewed as a sum of binary values, each indicating whether or not a particular model element is "on" or "off". Adopting the definition of area used in equation 3 leads to a joint objective function of the form,

$$\Phi(\mathbf{m}) \;=\; \phi_d \;+\; \lambda^2\phi_m \;=\; \|\mathbf{G}\,\mathbf{m}-\mathbf{d}\|_2^2 \;+\lambda^2\sum_{i=1}^{n}\frac{m_i^2}{m_i^2+\beta^2}, \tag{4}$$

where λ integrates both model element area and a secondary regularization parameter controlling the relative weighting of the two terms. Figure 1A shows the general form of the model regularization term (ϕ_m) in equation 4 for several values of β compared with the traditional case where $\mathbf{W} = \mathbf{I}$. The compactness term is clearly non-quadratic, though large values of β (relative to m_i) introduce a similar damping effect on the model parameters. For the case when $\beta << m_i$, the compactness constraint becomes apparent; the objective term asymptotes to one regardless of the magnitude of m_i. Hence, the penalty on model parameters does not depend on their relative magnitude, only whether or not they lie above or below the threshold of β. Minimization of this objective function yields a least-squares problem of the same form as equation 1 with the exception that the weighting matrix is now dependent on a model estimate,

$$\begin{bmatrix} \mathbf{G} \\ \lambda\mathbf{W}_c(\mathbf{m}^{j-1}) \end{bmatrix} \mathbf{m}^j \;=\; \begin{bmatrix} \mathbf{d} \\ 0 \end{bmatrix}, \tag{5}$$

where $\mathbf{W}_c(\mathbf{m})$ is a new diagonal matrix incorporating compactness and j is an index over sequential model estimates as discussed below. $\mathbf{W}_c$ can be written in explicit indicial form as,

$$W_{c_{ii}} \;=\left[m_i^2+\beta^2\right]^{-1/2}. \tag{6}$$

Since $\mathbf{W}_c$ is now dependent on $\mathbf{m}$, the resulting problem is non-linear and we must resort to iterative techniques, in this case a modified form of the iteratively reweighted least-squares (IRLS) method. Starting with a prior estimate of the model, equation 5 is solved for a new $\mathbf{m}$ followed by an update to the regularization operator. This solve/update sequence is repeated until a convergence criterion is met.

For the initial model estimate, we choose a smooth solution generated using standard regularization methods; experimentation suggests that the compactness result is relatively insensitive to the exact choice of starting model assuming that highly localized features are not already present. Figure 1B shows the linearized form of the model regularization term, ϕ_m, which now depends on the value of a prior model estimate, m_i^{j-1}. First, note that even for small values of β (relative to m_i^{j-1}) the objective function does not asymptote to a constant value as in figure 1A. In this case where $\beta << m_i^{j-1}$, it is the ratio m_i^j/m_i^{j-1} that is penalized quadratically. When $\beta >> m_i^{j-1}$, the objective term simplifies to the traditional damping case and β serves to bound the maximum value of any element on the diagonal of $\mathbf{W}_c$; as $m_i \to 0$, $W_{ii} \to 1/\beta$. At any given step, $\mathbf{W}_c$ should be viewed as a spatially variable damping matrix with high values in regions where the prior model estimate has a small absolute magnitude. This promotes compactness and reduces unnecessary model complexity by damping out relatively small amplitude features within the constraints provided by the data misfit term.

1.2 Misfit Levels, (λ, β) Selection, and Convergence Tests

The algorithm as outlined above neglects two issues, the selection of the regularization parameters (λ, β) and when to terminate the IRLS process. As can be seen in the definition of the least squares problem (equation 5), both λ and β affect the strength of the model weighting operator in an interconnected fashion; β bounds the maximum value in $\mathbf{W}_c$ while λ scales all of the elements. Small values of β correspond to the limiting case of interest in equation 3, but introduce instability as $m_i \to 0$. Larger values of β provide a more stable result at the cost of additional smoothness in the model. λ behaves like a more traditional regularization parameter but scales a spatially variable damping matrix which changes at each IRLS iteration.

The solution generated by an optimal choice of parameters should both fit the data within some tolerance and exhibit structural features consistent with prior information. However, care must be taken not

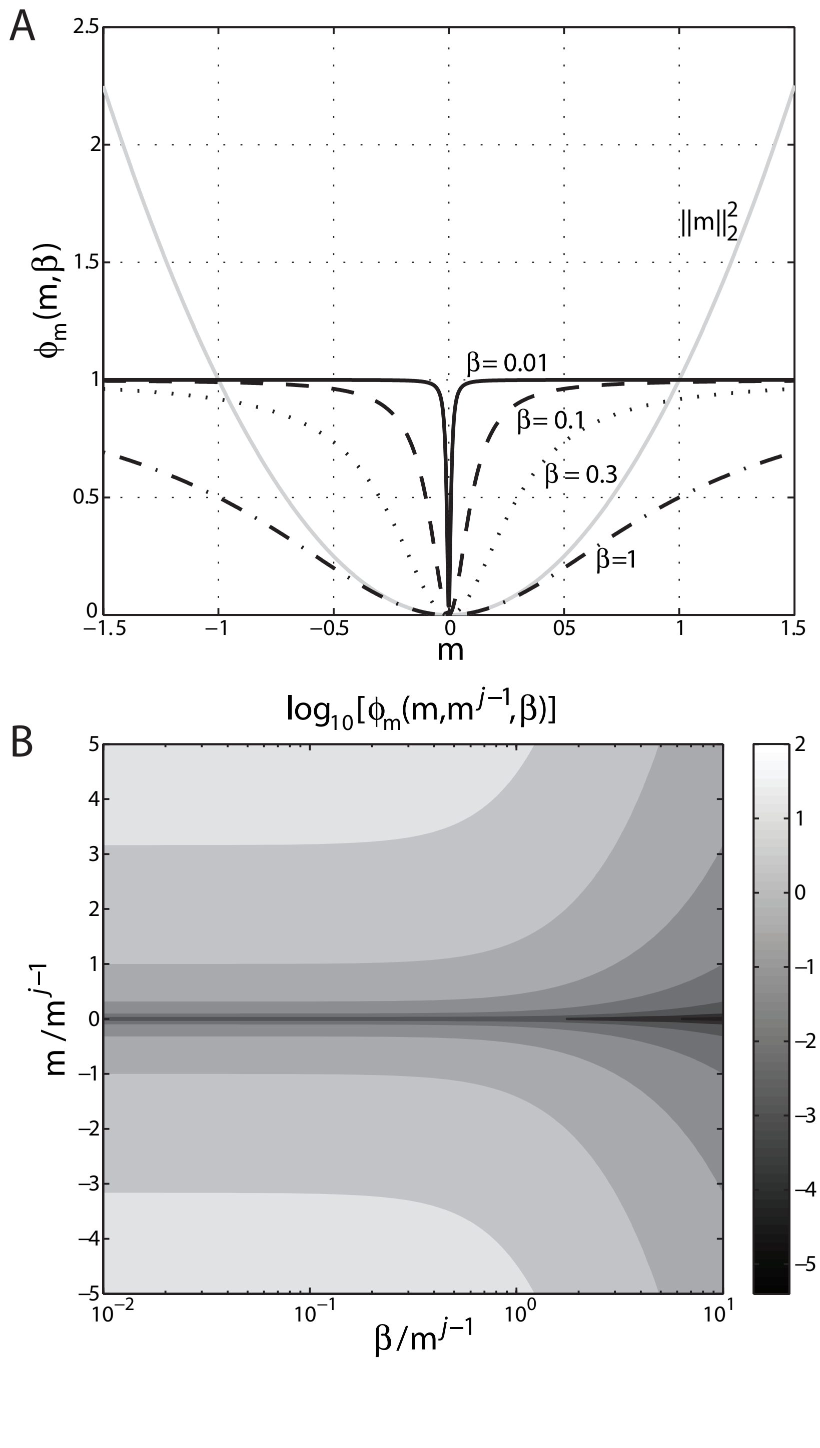

Figure 1: (A) Illustration of the non-quadratic term in the objective function for the compactness constraint defined in equation 4 for several values of β. These are compared with a traditional quadratic term (grey line). (B) The linearized version of the compactness term in the objective function, which is now a function of both β and a prior model estimate $\mathbf{m}^{j-1}$.

to *overfit* the data; if measurement error exists, a perfect fit will map noise into model structure. Morozov's discrepancy principle [*Morozov*, 1962] suggests that the best regularization level is the one which yields the "simplest" model which still satisfies the data to within a bound determined by data variance, $\|\mathbf{G}\,\mathbf{m} - \mathbf{d}\|_2 \leq \epsilon$. Unfortunately, *a priori* estimates of data variance are not available for most geophysical measurements making direct application of the discrepancy principle difficult. Additionally, most problems involving multiple regularization components have a range of parameter pairs which produce solutions with both the correct data residual and equivalent model norms. Selection from within this class of simple, equally feasible solutions must rely on prior knowledge, which is typically a judgment made by the geophysicist.

The L-curve method [*Hansen*, 1992; *Hansen and O'Leary*, 1993] and the generalized cross validation (GCV) method [*Wahba*, 1977, 1990] are two approaches for selecting regularization parameters which do not require noise estimates. The L-curve method involves calculation of a trade-off curve between data residual and model semi-norm for a selection of regularization parameter values; the value corresponding to the point of maximum curvature on this curve is considered optimal. The GCV method selects a regularization parameter which maximizes solution robustness by evaluating the stability of model estimates when selected data points are eliminated. This technique requires finding the parameter value associated with the minimum of this GCV cost function. Although both methods have been successfully applied to geophysical inverse problems, they are sometimes unstable in cases where correlated noise is present. More crucially, they suffer from the same problem as the discrepancy method when considering multiple regularization parameters e.g. 2 dimensional GCV surfaces often exhibit minima resembling troughs, thus necessitating the addition of prior knowledge to select an appropriate parameter pair.

Previous investigations of compactness constraints have used a combination of the above techniques, experimentation, and interpretive skills to determine the correct (λ, β) pairs. The original work of *Last and Kubik* [1983] suggests that β should have a value close to machine precision ($\approx 10^{-11}$ in their case) with λ determined using the discrepancy principle and an assumed, although somewhat arbitrary, noise level. *Zhdanov and Tolstaya* [2004] advocates the use of a procedure similar to the L-curve technique where β is chosen to be the point of maximum curvature on the trade-off curve relating β to $A(\mathbf{m})$. We have observed that setting β to values near machine precision results in severe instability and that the approach of *Zhdanov and Tolstaya* [2004] often yields trade-off curves with poorly defined corners. We therefore fix β at a reasonable value determined by experience, typically between 10^{-4} and 10^{-7}. This reduces the regularization parameter search problem to a single dimension. We then select an appropriate λ value by using the L-curve approach, adding manual guidance if the curve is not well-behaved. λ and β are kept fixed through later IRLS iterations although dynamic re-adjustment of λ might be a superior approach [*Farquharson and Oldenburg*, 2004]. In cases where compactness results are compared to other inversion methods, we first determine an optimal λ for the Tikhonov problem and then use the resulting data misfit value to determine λ for the compactness inversion. This approach guarantees that the inversion results for both methods are equiprobable from a data fit perspective.

A second concern is the formulation of a stopping criterion for the IRLS procedure. In our implementation we use a bound (α) on the change in the area metric between non-linear iterations to terminate the procedure, $\|A(\mathbf{m}^j) - A(\mathbf{m}^{j-1})\| \leq \alpha$. In some cases, the process is halted manually if the solution matches a prior conception of structure. Alternative stopping criterion formulated in terms of data misfit reduction are problematic because we find that, after the first iteration, the l_2 data residuals are almost constant as a function of IRLS iteration.

1.3 Compactness And Traveltime Tomography

Our discussion of compactness so far has been general with no assumptions regarding the operation which $\mathbf{G}$ performs, the model parametrization represented by $\mathbf{m}$, or the type of data stored as $\mathbf{d}$. We will now apply our formulation to the concrete example of differential seismic traveltime tomography. In the case of differential inversion we consider two datasets, $\mathbf{d_1}$ and $\mathbf{d_2}$, acquired at different times, but with the same geometry. Instead of inverting the two datasets independently, we invert $\Delta\mathbf{d}$, the difference between measurements made at time 1 and time 2, for $\Delta\mathbf{m}$, the temporal perturbation in model parameters. The starting system becomes $\mathbf{G}\,\Delta\mathbf{m} = \Delta\mathbf{d}$ and compactness constraints are applied directly to the model perturbations. When

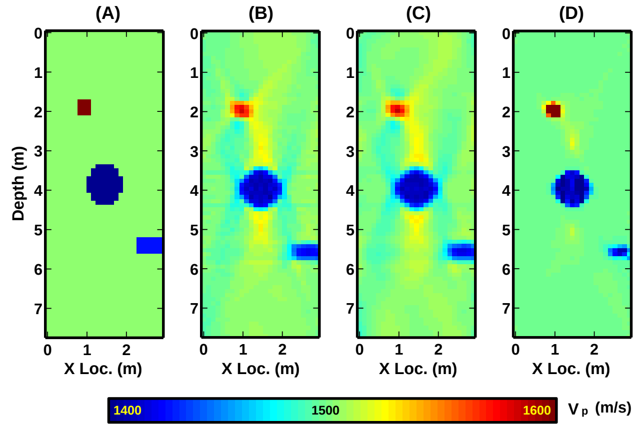

Figure 2: A noise-free synthetic test : (A) True velocity model, (B) 1st order Tikh. regularization, (C) 2nd order Tikh. regularization, (D) Compactness constraints.

applying this approach to traveltime tomography, we choose $\Delta\mathbf{m}$ to be a 2D rectilinear mesh of constant slowness cells while $\Delta\mathbf{d}$ is a vector of differenced first-arrival traveltimes and $\mathbf{G}$ is the ray-path matrix.

At each iterative step in the inversion, a coupled system of the form shown in equation 5 is solved using the LSQR algorithm [*Paige and Saunders*, 1982]. The starting model required to compute the first weighting matrix is calculated using the same $\mathbf{G}$, but with 1st order Tikhonov regularization instead of $\mathbf{W}_c$. This procedure minimizes the area metric with respect to perturbations from the background estimate.

We restrict our consideration to examples where $\mathbf{G}$ is linear, which corresponds to situations where ray-paths are fixed within the inversion process. In differential tomography this is often a reasonable assumption since timelapse perturbations are typically small in comparison to background property variations, which allows the use of rays traced in a prior reference model for the inversion. In situations where the assumption of a linear operator are clearly invalid, the IRLS process can be extended to include variable ray-paths by updating $\mathbf{G}$ at each iteration. We leave examination of the more complete non-linear problem including variable ray-curvature to future investigations and focus on the case where the only non-linearity present is introduced via the regularization operator.

2 A Simple Synthetic Test

The compactness algorithm we describe was initially tested on two static synthetic crosswell datasets to enhance our understanding of the IRLS process, corresponding changes in the weighting matrix, and the role of the compact body assumption. Straight ray traveltimes were first generated for a symmetric 40×40 source/receiver configuration (1600 data) using the velocity model shown in panel [A] of figure 2 with three compact perturbations. For the inversion, model estimates were calculated on a 25×75 sample mesh. All inversions used identical versions of the modeling operator, $\mathbf{G}$, and differ only in constraint implementation.

The right three panels of figure 2 depict noise-free inversion results using both standard 1st and 2nd order Tikhonov regularization (panels [B] and [C]) and compactness constraints (panel [D]). In all examples, a β value of 10^{-5} was used for the compactness inversions. A target data residual was chosen by interpretive analysis of the 1st order Tikhhonov results guided by L-curve analysis. λ values were selected for each inversion to match this misfit level, and the results for all three methods fit the data equally well. Object smearing, due to limited angular aperture, is visible in both standard tomograms. Artifacts of this type often plague traveltime imaging results and obscure both qualitative interpretation and the recovery of quantitative property estimates. As can be seen in panel [D], the addition of compactness constraints largely eliminates object smearing. The compactness inversion was initialized using an over-damped solution with 1st order

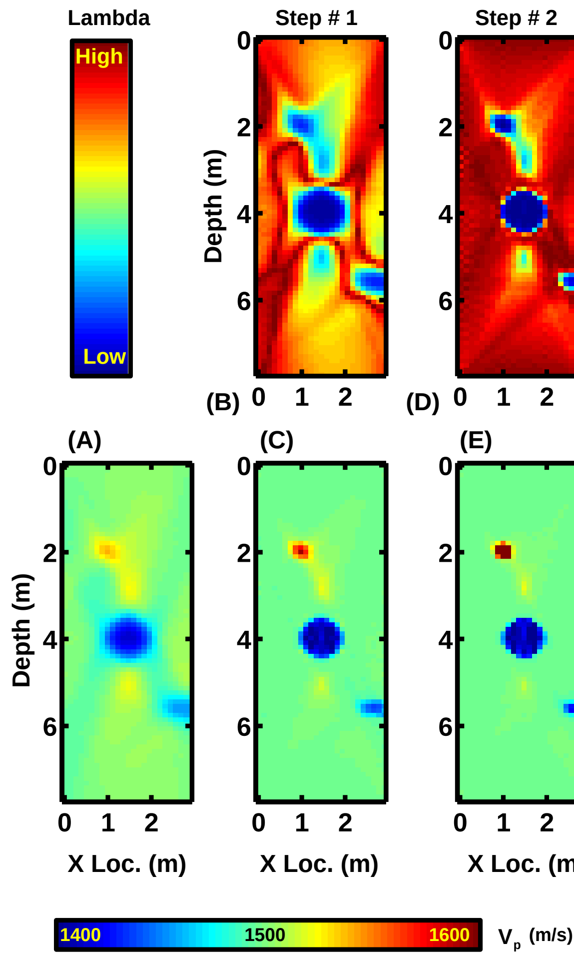

Figure 3: $\mathbf{W}_c$ and $\mathbf{m}$ as a function of iteration: The top row depicts $diag(\mathbf{W}_c)$, the spatial variations in damping. The bottom row shows the corresponding estimates of $\mathbf{m}$. Panel (A) shows the starting model, generated using 1st order Tikh. constraints

Tikhonov regularization. In this case, the IRLS loop converged in only 2 IRLS iterations.

During the IRLS procedure, $\mathbf{W}_c(\mathbf{m})$ changes in accordance with variations in the previous model estimate. Examining the spatial characteristics of $\mathbf{W}_c$ provides insight into the way which the compactness constraints evolve. Figure 3 shows images of $diag[\mathbf{W}_c]$ (top row) and the corresponding estimates of $\mathbf{m}$ (bottom row) for the first two IRLS iterations. The first model estimate [A] shows the smooth over-damped starting model. The resulting $\mathbf{W}_c$ operator [B] exhibits large damping values in zones with no anomalous features and smaller values in the vicinity of the three perturbations. The second and third iterations exhibit increasingly tight constraints around the perturbed zone, which corresponds to a reduction in perturbation area, $A(m)$. After the 1st iteration, the l_2 data residual is essentially constant as a function of IRLS iteration for this choice of β and λ. Interestingly, some of the artifacts present in the starting model are still visible in later iterations except in a focused form.

Figure 4 depicts results from the same synthetic problem with the addition of 3% gaussian noise to the data. As in the noise-free case, compactness constraints minimize artifacts due to limitations in survey aperture. Additionally, image artifacts due to noise in the traveltime data are partially suppressed yielding

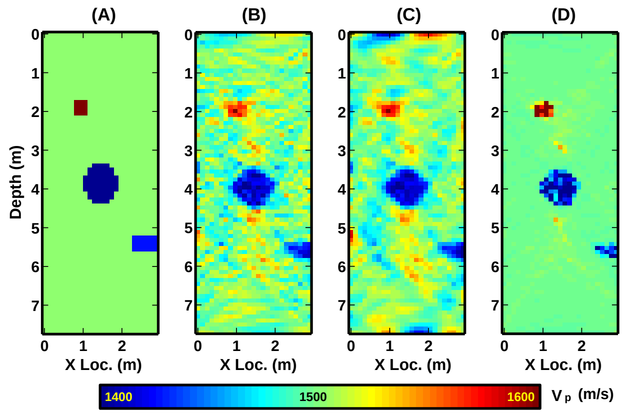

Figure 4: A synthetic test with 3% gaussian noise: (A) True velocity model, (B) 1st order Tikh. regularization, (C) 2nd order Tikh. regularization, (D) Compactness constraints.

a more interpretable image. However, the compact solution exhibits some less desirable features including a reduction in the size of the lower right velocity anomaly. Imaging artifacts above and below the central anomaly are also focused into small high amplitude features. This illustrates an important aspect of the compactness constraint; model resolution is data dependent due to the introduction of a regularization operator that depends on a prior model estimate. Therefore, the extent to which a truly compact feature can be recovered depends on the data quality.

As alluded to previously, choosing a compact model is one of many techniques for introducing prior information to the inversion process. Although we believe that compact solutions are appropriate for imaging flow-induced features, diffusive processes such as thermal conduction might result in smooth velocity anomalies. In such cases, regularization methods favoring smooth solutions will yield superior results. Figure 5 depicts a comparison of Tikhonov and compactness solutions for a case where the true model is smooth. The true model (panel [A]) was generated by applying a gaussian filter to the model shown in panel [A] of figure 2, followed by a re-normalization to yield the same maximum and minimum velocities. Synthetic data were generated using the same geometry and noise levels used in the previous example.
As can be seen in figure 5, both the 1st order (panel [B]) and 2nd order (panel [C]) Tikhonov solutions effectively recover the true model. The addition of compactness constraints (panel [D]) yields a model estimate with a grainy texture and locally compact features. While correctly locating the velocity anomalies, the compact solution does not effectively capture the smooth transitions around the velocity features and underestimates their spatial extent. This pair of examples clearly demonstrates the benefit of incorporating prior knowledge into the regularization process. In this case, constraints based on a valid understanding of anomaly structure significantly improve image quality, motivating their inclusion when such insight is available.

3 A Synthetic Contaminant Infiltration Example

The previous examples considered only a set of simplified features with no particular physical significance. Evaluating the efficacy of the algorithm for fluid process monitoring applications requires a synthetic test case where the imaging target exhibits the geometry of a flow-induced property perturbation. The geophysical monitoring of dense non-aqueous phase liquids (DNAPLs) infiltrating the subsurface is one of many imaging problems which exhibits these spatial characteristics. Because DNAPLs are denser and often less viscous than water, they easily penetrate deep into the saturated zone and pond at low permeability barriers [*Pankow and Cherry*, 1996]. GPR imaging results from the Borden field experiment [*Greenhouse et al.*, 1993] [*Brewster*

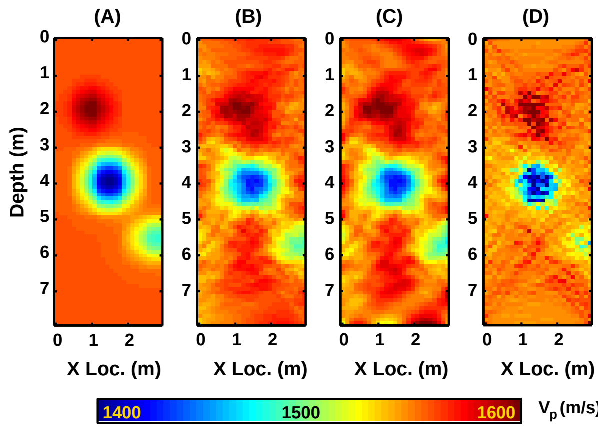

Figure 5: A smooth synthetic test with 3% gaussian noise: (A) True velocity model, (B) 1st order Tikh. regularization, (C) 2nd order Tikh. regularization, (D) Compactness constraints.

and Annan, 1994] revealed that DNAPLs injected at the site in question formed several thin lenticular zones of high saturation, visible as laterally discrete reflection events. *Parker et al.* [2003] used direct-push fluid sampling techniques to quantify the vertical distribution of DNAPLs at five contaminated sites; this study confirmed that dense contaminants often form very thin (5 to 15 cm thick) pools of limited lateral extent, features which we would characterize as "compact". Motivated by these observations, we selected the 2-phase contaminant modeling results of *Kueper and Gerhard* [1995] as the geometric basis for our second synthetic test. In this case, we synthesize and invert differential crosswell radar traveltimes. Changes in bulk dielectric constant, and therefore radar velocities, are relatively sensitive to DNAPL saturation as shown in laboratory investigations [*Ajo-Franklin et al.*, 2004], making GPR the preferred imaging modality.

The study of *Kueper and Gerhard* [1995] examined the infiltration characteristics of point-source spills of tetrachloroethylene (PCE) for spatially correlated random permeability models. The models used for their numerical experiments were 50.5×20.125 m two dimensional permeability fields with anisotropic exponential autocorrelation functions with correlation lengths of 5.0 m and 0.5 m in the horizontal and vertical directions respectively. They used an adaptation of the two-phase flow code presented in *Kueper and Frind* [1991] to simulate the release of 5.0 m^3 of PCE (ρ_{fl} = 1460 kg/m^3, η = 0.00057 Pa s). We use several snap-shots from one of their flow realizations as starting points for our numerical experiments. Panel (A) of figure 6 shows the results for one of their flow simulations with a background image of log permeability structure. The blue, green, and red zones correspond to regions with greater than 20 % PCE saturation at three times within the simulation.

We started with the stochastic permeability model and multiphase saturation results detailed in *Kueper and Gerhard* [1995] and adapted them to provide the requisite input parameters for radar modeling and tomography. The log of the initial permeability models were linearly mapped to porosities ranging from 0.3 to 0.48 with more hydraulically conductive regions mapped to higher porosities. The DNAPL saturation results and the porosity values were converted to maps of dielectric constant using the Bruggeman/Hanai/Sen (BHS) model [*Sen et al.*, 1981]. We assumed a 3-phase system composed of quartz ($\kappa = 4.27$, *Keller* [1987]), water ($\kappa = 79.8$, *Sen et al.* [1981], and PCE ($\kappa = 2.297$, *Nath and Narain* [1982]) for dielectric property estimation. Panel (B) of figure 6 shows the estimated dielectric constant model for time C. We also assumed that the imaginary component of the dielectric constant was 0 to facilitate conversion to a radar slowness model. Using the slowness map and the source/receiver geometry as input, synthetic radar traveltimes were generated using straight-ray tracing on a 90 x 190 sample mesh with $\Delta x = \Delta z = 0.1$ m. First arrival times for the plume at times B and C were subtracted to yield a Δt dataset suitable for difference inversion. Gaussian noise (2.5 %) was added to the time differenced dataset.

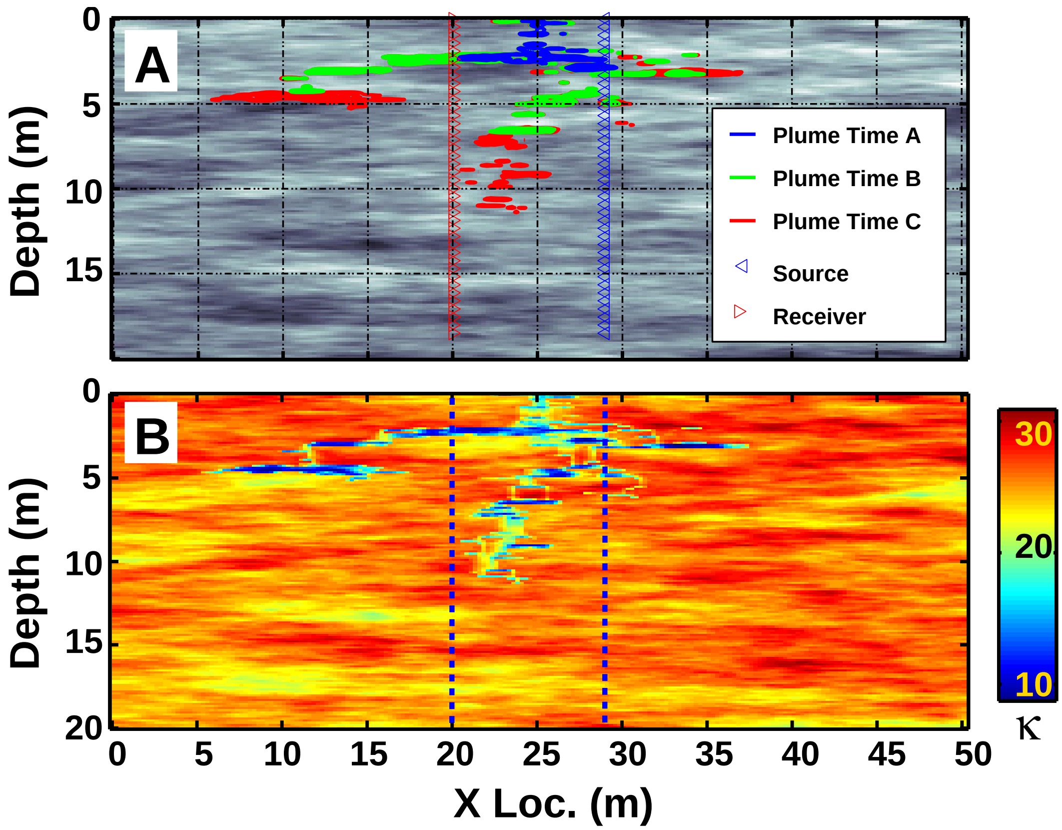

Figure 6: DNAPL pool geometry and dielectric model : Panel (A) : The background model permeability variations are shown in greyscale while the evolution of pool geometry at three successive times is shown in solid colors. The source (blue triangles) and receiver (red triangles) locations for the synthetic crosswell experiment are also shown. Panel (B) shows the dielectric model at time C. Dielectric values were estimated from the flow results and the BHS model. The dashed blue lines indicate the extent of the synthetic crosswell dataset.

Figure 7 shows the tomographic result generated by both traditional 1st and 2nd order Tikhonov regularization and the inversion with compactness constraints. All inversions were performed on 45 x 95 sample rectilinear meshes with $\Delta x = \Delta z = 0.2$ m. All slowness models shown use the same colorscale to allow for accurate comparisons. Panel [A] shows the true differential slowness model, $\Delta s = s_C - s_B$. The geometry of the zone with DNAPL induced slowness changes is composed of a series of vertically connected lenticular features with variable saturation levels. Panels [B] through [E] show the tomography results using 1st (left) and 2nd (right) order Tikhonov regularization with both small (bottom) and large (top) values for λ. Panel [F] shows the result of the inversion based on compactness constraints after four reweighting steps. Like the results from the first synthetic test, the only difference between the inversions are the constraint implementations.

Several conclusions can be drawn from figure 7. When considering the standard Tikhonov solutions, small λ values allow more detailed delineation of the DNAPL zones at the cost of increased image noise levels. Larger λ values succeed in suppressing artifacts but simultaneously smear the target feature. In contrast, the application of compactness constraints succeeds in both decreasing artifacts and providing a sharper image of the DNAPL zone. The compactness based inversion also provides more accurate slowness estimate, largely because the anomaly is focused into an appropriate geometry.

4 Analysis Of The Frio CO_2 Monitoring Dataset

For our final algorithmic test we explored the use of differential seismic traveltime tomography with compactness constraints to monitor the subsurface movement of supercritical carbon dioxide (CO_2). The injection of carbon dioxide into deep saline aquifers is one approach currently being considered to minimize anthropogenic contributions to atmospheric greenhouse gasses [*Orr*, 2003]. Since the goal of sequestration is the long-term storage of CO_2, remote monitoring is required to ensure that reservoir seals remain intact [*Benson*, 2003]. Seismic methods offer one possible approach to mapping the subsurface extent of CO_2 saturation, a topic explored in previous enhanced oil recovery (EOR) monitoring studies [*Lazaratos and Marion*, 1997] [*Davis et al.*, 2003]. Borehole imaging methods in particular have shown promise for resolving small-scale flow features in a CO_2 monitoring context [*Majer et al.*, 2006].

A complete test of the compactness algorithm was performed on a timelapse crosswell monitoring dataset acquired at the Frio pilot site. The Frio demonstration project [*Hovorka et al.*, 2006] is an on-going multi-institution effort to improve understanding of the *in situ* dynamics of CO_2 injection within a saline aquifer located in East Texas. In the first stage of the project, 1600 $\times 10^3$ kg of supercritical CO_2 (at $P = 15$ MPa, $T = 55^o$C) was injected into a confined unit of the Oligocene Frio sandstone formation at a depth of 1534 m (5053 ft). The target unit, referred to as the Frio "C" sand, has a dip of approximately 15 degrees. Cores from the "C" sand exhibited porosities between 30 and 35 % with permeabilities between 2000 and 2500 md. *Hovorka et al.* [2006] includes a complete description of the formation properties and the multiple facets of the Frio experiment. Since the supercritical phase at these P/T conditions is characterized by both a low density (≈ 700 kg/m^3, *Hovorka et al.* [2006]) and a low bulk modulus (≈ 0.086 GPa, *Wang and Nur* [1989]), the zone of CO_2 saturation was expected to migrate up-dip and be visible as a zone of decreased P-wave velocity.

Motivated by previous seismic monitoring projects [*Lazaratos and Marion*, 1997], Lawrence Berkeley National Laboratory acquired two high-quality crosswell seismic datasets before and after the pilot injection in an attempt to delineate the region of subsurface CO_2 saturation. *Daley et al.* [2005] describes the data collection procedure and relevant survey parameters. Panel A of figure 8 shows a schematic representation of the lithological units within the survey domain including the "C" sand used for the injection experiment. Traveltimes from the baseline and repeat surveys were picked and subtracted to yield time differences (Δt). This dataset, consisting of 3301 differential times, was then inverted to generate a map of changes in slowness (Δs). Δs images were converted to maps of velocity changes using a background reference model obtained from logs and the baseline survey. Figure 8 shows tomography results for the Frio dataset using 0th [panel B], 1st [panel C], and 2nd [panel D] order Tikhonov regularization in comparison to compactness constraints [panel E]. All inversions were performed on a 60×200 sample rectilinear mesh.

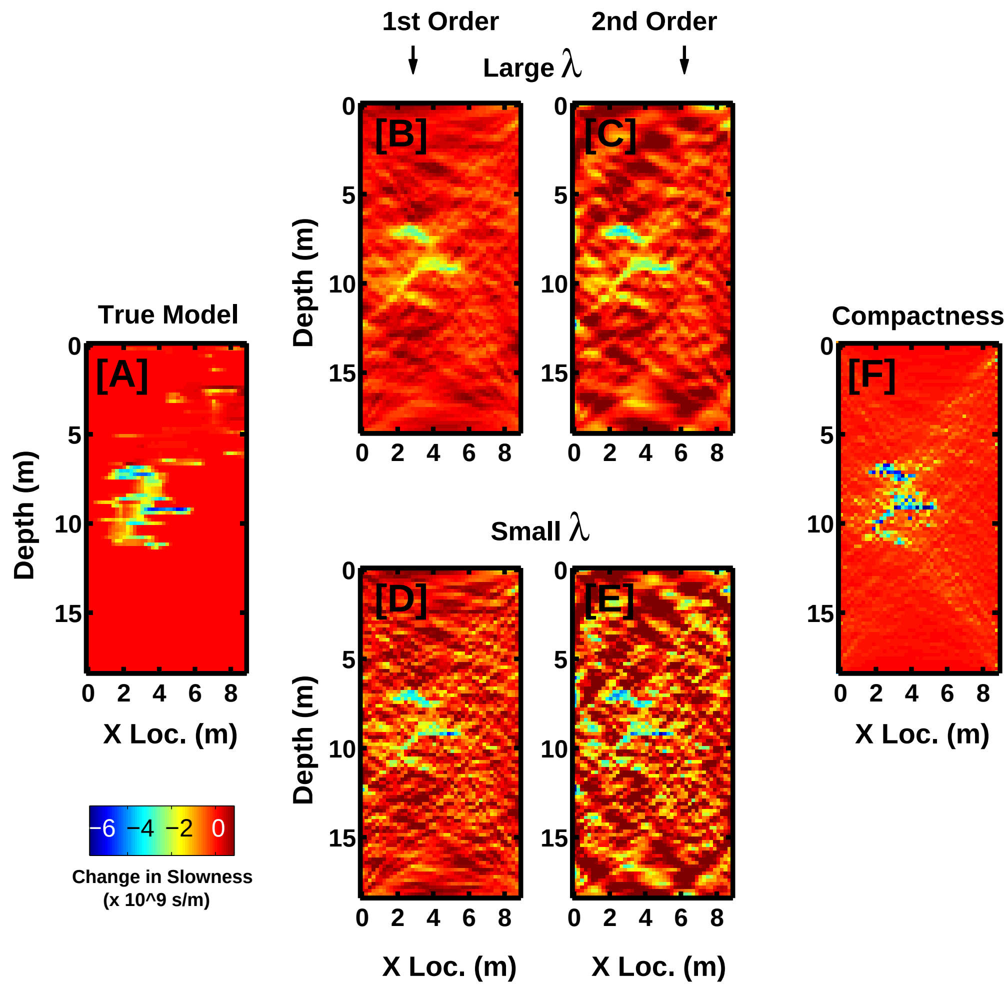

Figure 7: Tomography results for the DNAPL test case : Panel [A] shows the true differential slowness model. Panels [B] and [C] show difference tomography results using 1st and 2nd order Tikhonov regularization with large λ values. Panels [D] and [E] show equivalent results calculated using weaker constraints. Panel [F] shows the optimal result using compactness constraints.

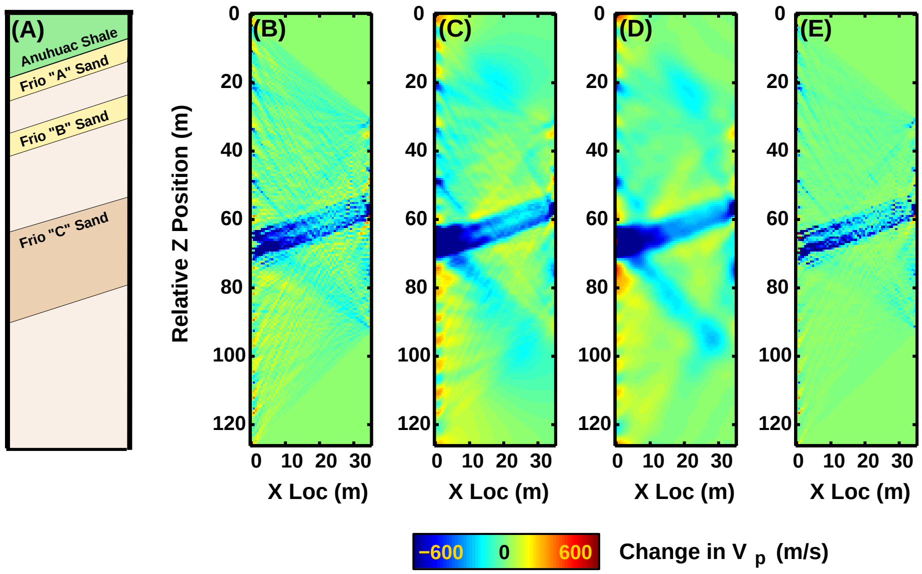

Figure 8: Results from the Frio crosswell monitoring experiment: (A) Schematic representation of units within the survey domain, differential tomograms with 0th order Tikh. regularization (damping) (B), 1st order Tikh. (C), 2nd order Tikh. (D), and compactness constraints (E). The coordinate system used in panels (B), (C), (D), and (E) is relative to the uppermost measurement source location.

The region of injected CO_2 is visible in all four images as a linear feature with decreased P-wave velocity. The dip of the imaged CO_2 zone matches prior models of local structure, increasing our confidence in the reconstruction. As can be seen in figure 8, the results based on 1st and 2nd order Tikhonov schemes (panel C and D) were generally of low quality with significant artifacts including a broad low velocity feature which extends counter to structural orientation and ray related features near the boreholes. Both the 0th order Tikhonov and compactness constrained models exhibit a high degree of similarity within the anomalous zone. However, the addition of compactness constraints has largely eliminated artifacts related to ray coverage; these artifacts are particularly visible on the left side of the 0th order solution. Since ensuring seal integrity is one of the primary goals of sequestration monitoring, a robust interpretation requires the reduction of low velocity artifacts outside of the formation.

The maximum magnitude of the velocity perturbation observed in our reconstruction (≈ -600 m/s) is surprisingly large. The neglect of ray curvature in the reconstruction is another possible source of error; however the inclusion of ray bending typically increases rather than decreases the magnitude of velocity anomalies. Preliminary rock-physics analysis using poroelastic fluid substitution models similar to those used in *Nolen-Hoeksema et al.* [1995] indicate that such effects are insufficient to produce this change. Likewise, *Wang et al.* [1998] observe significantly smaller reductions in V_p during core-scale laboratory measurements of CO_2 injection. The absence of significant changes in V_s (not shown) suggests that an increase in pore pressure can also be ruled out as a secondary factor. Future investigation is needed to identify the mechanism responsible for these large changes in rock properties.

5 Conclusion

We describe one approach for including compactness constraints in differential seismic traveltime tomography through use of a model-space reweighting algorithm. We observe improvements in image quality in comparison to standard Tikhonov-based techniques for a simple test problem, a synthetic based on multiphase flow results, and a CO_2 sequestration monitoring dataset. For the test problems including localized features, the addition of compactness regularization reduces tomographic artifacts, improves the recovery of target geometry, and more quantitatively estimates property variations. The strength of compactness constraints can also be a weakness; in situations where the imaging target exhibits smooth variations, the size of the reconstructed feature can be inappropriately reduced, potentially resulting in property overestimates. The assumption of compactness is particularly dangerous when considering flow processes dominated by diffusion rather than advection. One promising extension to our approach would be to replace compactness with connectivity as a secondary constraint in the reweighting process, possibly a more effective regularization approach for flow imaging. A general limitation of differential inversion is the requirement that all surveys have matching geometries, a requirement which is often difficult to fulfill in practice. A possible solution to this limitation would be to pose the imaging problem as a coupled joint inversion across multiple time steps [*Day-Lewis et al.*, 2003], thus allowing the inclusion of compactness constraints on model perturbations without requiring explicit data differencing.

6 Acknowledgments

The DNAPL flow modeling results presented were generously provided by J.I. Gerhard. and B.H. Kueper of Queens University. We would like to thank Prof. M.N. Toksöz and the Founding Members Consortium of the Earth Resources Laboratory for their generous support and guidance. Lawrence Berkeley National Laboratory collected the Frio seismic datasets with support from the U.S. Dept. of Energy's GEOSEQ program. We would also like to thank Sally Benson, Ernie Majer, Larry Myer, Susan Hovorka, the Texas Bureau of Economic Geology, and all of the contributors to the Frio project.

References

Acar, R., and C. Vogel (1994), Analysis of bounded variation penalty methods for ill-posed problems, *Inverse Problems*, *10*(6), 1217–1229.

Ajo-Franklin, J., J. Geller, and J. Harris (2004), The dielectric properties of granular media saturated with DNAPL/water mixtures, *Geophysical Research Letters*, *31*(L17501).

Aster, R., B. Borchers, and C. Thurber (2005), *Parameter Estimation and Inverse Problems*, Elsevier Academic Press.

Benson, S. (2003), Geologic sequestration of carbon dioxide, in *The Carbon Dioxide Dilemma : Promising Technologies and Policies*, National Academies Press.

Bertete-Aguirre, H., E. Cherkaev, and M. Oristaglio (2002), Non-smooth gravity problem with total variation penalization functional, *Geophysical Journal International*, *149*(2), 499–507.

Brewster, M., and A. Annan (1994), Ground-penetrating radar monitoring of a controlled DNAPL release: 200 MHz radar, *Geophysics*, *59*(8), 1211–1221.

Bube, K., and R. Langan (1997), Hybrid l^1/l^2 minimization with applications to tomography, *Geophysics*, *62*, 1183–1195.

Claerbout, J., and F. Muir (1979), Robust modeling with erratic data, *Geophysics*, *38*(5), 826–844.

Constable, S., R. Parker, and C. Constable (1987), Occam's inversion: A practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, *52*(3), 289–300.

Daley, T., L. Myer, and E. Majer (2005), Acquisition of time-lapse, 6-component, P- and S-wave, crosswell seismic survey with an orbital vibrator and of time-lapse VSP for CO_2 injection monitoring, in *75th Ann. Internat. Mtg*, Soc.Of Expl. Geophysocs.

Davis, T., M. Terrel, R. B. R. Cardona, R. Kendall, and R. Winarsky (2003), Multicomponent seismic characterization and monitoring of the CO_2 flood at Weyburn Field, Saskatchewan, *Leading Edge*, *22*, 696–697.

Day-Lewis, F., J. Lane, J. Harris, and S. Gorelick (2003), Time-lapse imaging of saline tracer tests using cross-borehole radar tomography, *Water Resources Research*, *39*(10), doi:10.1029/2002WR001722.

Farquharson, C., and D. Oldenburg (2004), A comparison of automatoc techniques for estimating the regularization parameter in non-linear inverse problems, *Geophysical Journal International*, *156*, 411–425.

Farquharson, C. G., and D. W. Oldenburg (1998), Non-linear inversion using general measures of data misfit and model structure, *Geophysical Journal International*, *134*, 213–227.

Greaves, R., and T. Fulp (1987), Three-dimensional seismic monitoring of an enhanced oil recovery process, *Geophysics*, *52*(9), 1175–1186.

Greenhouse, J., M. Brewster, G. Schneider, . D. Redman, P. Annan, G. Olhoeft, J. L. K. Sander, and A. Mazella (1993), Geophysics and solvents : The Borden experiment, *The Leading Edge*, *12*(4), 261–267.

Guillen, A., and V. Menichetti (1984), Gravity and magnetic inversion with minimization of a specific functional, *Geophysics*, *49*(8), 1354–1360.

Hansen, P. (1992), Analysis of discrete ill-posed problems by means of the L-curve, *SIAM Review*, *34*(4), 561–580.

Hansen, P., and D. O'Leary (1993), The use of the L-curve in the regularization of discrete ill-posed problems, *SIAM Journal of Scientific Computing*, *14*, 1487–1503.

Hovorka, S., et al. (2006), Measuring permanence of CO_2 storage in saline formations : the Frio experiment, *Environmental Geosciences*, *13*(2), 1–17.

Keller, G. (1987), Rock and mineral properties, in *Electromagnetic Methods In Applied Geophysics*, edited by M. N. Nabighian, pp. 13–51, Society Of Exploration Geophysics, Tulsa, Oklahoma.

Kueper, B., and E. Frind (1991), Two-phase flow in heterogeneous porous media 1. Model development, *Water Resources Research*, *27*(6), 1049–1057.

Kueper, B., and J. Gerhard (1995), Variability of point source infiltration rates for two-phase flow in heterogeneous porous media, *Water Resources Research*, *31*(12), 2971–2980.

Kueper, B., D. Redman, R. Starr, S. Reitsma, and M. Mah (1993), A field experiment to study the behavior of tetrachloroethylene below the water table : Spatial distribution of residual and pooled DNAPL, *Ground Water*, *31*(5), 756–766.

Last, B., and K. Kubik (1983), Compact gravity inversion, *Geophysics*, *48*(6), 713–721.

Lazaratos, S., and B. Marion (1997), Crosswell seismic imaging of reservoir changes caused by CO_2 injection, *The Leading Edge*, *16*, 1300–1306.

Majer, E., T. Daley, V. Korneev, D. Cox, J. Peterson, and . J. Queen (2006), Cost-effective imaging of CO_2 injection with borehole seismic methods, *The Leading Edge*, *25*(10), 1290–1302.

Morozov, V. (1962), The choice of parameter when solving functional equations by regularization, *Dokl. Akad. Nauk SSSR*, *175*, 1225–1228.

Nath, J., and B. Narain (1982), Binary systems of tetrachlorethylene with benzene, toluence, p-xylene, carbon tetrachloride, and cyclohexane. 1. Ultrasonic velocities and adiabatic compressibilities at 293.15 and 303.15 K, dielectric constants at 298.15 and 308.15 K, and refractive indexes at 298.15 K, *J. Chem. Eng. Data*, *27*, 308–312.

Nolen-Hoeksema, R., Z. Wang, J. M. Harris, and R. T. Langan (1995), High-resolution crosswell imaging of a west Texas carbonate reservoir: Part 5: Core analysis, *Geophysics*, *60*(3), 712–726.

Orr, F. M. (2003), Sequestration via injection of carbon dioxide into the deep earth, in *The Carbon Dioxide Dilemma : Promising Technologies and Policies*, National Academies Press.

Paige, C., and M. Saunders (1982), An algorithm for sparse linear equations and sparse least squares, *ACM Transactions in Mathematical Software*, *8*, 43–71.

Pankow, J., and J. Cherry (1996), *Dense Chlorinated Solvents and other DNAPLs in Groundwater*, Waterloo Press.

Parker, B., J. Cherry, S. Chapman, and M. Guilbeault (2003), Review and analysis of chlorinated solvent dense nonaqueous phase liquid distributions in five sandy aquifers, *Vadose Zone Journal*, *2*, 117–137.

Portniaguine, O., and M. Zhdanov (1999), Focusing geophysical inversion images, *Geophysics*, *64*(3), 874–887.

Rudin, L., S. Osher, and E. Fatemi (1992), Nonlinear total variation based noise removal algorithms, *Physica D*, *60*(1-4), 259–268.

Scales, J., A. Gersztenkorn, and S. Treitel (1988), Fast l^p solution of large, sparse, linear-systems: Application to seismic travel time tomography, *Journal of Computational Physics*, *75*(2), 314–333.

Sen, P., C. Scala, and M. Cohen (1981), A self-similar model for sedimentary rocks with application to the dielectric constant of fused glass beads, *Geophysics*, *46*(5), 781–795.

Tikhonov, A., and V. Arsenin (1977), *Solutions of Ill-posed Problems*, John-Wiley & Sons, New York, NY.

Wahba, G. (1977), Practical approximate solutions to linear operator equations when the data are noisy, *SIAM Journal on Numerical Analysis*, *14*(4), 651–667.

Wahba, G. (1990), *Spline Models for Observational Data*, *CBMS-NSF Regional Conference Series in Applied Mathematics*, vol. 59, Society For Industrial And Applied Mathematics.

Wang, Z., and A. Nur (1989), Effects of CO_2 flooding on wave velocities in rocks with hydrocarbons, *SPE Reservoir Engineering*, *4*(4), 429–436.

Wang, Z., M. Cates, and R. Langan (1998), Seismic monitoring of a CO_2 flood in a carbonate reservoir : A rock physics study, *Geophysics*, *63*(5), 1604–1617.

Youzwishen, C., and M. Sacchi (2006), Edge preserving imaging, *Journal of Seismic Exploration*, *15*(15), 45–57.

Yu, M., and D. Dougherty (2000), Modified total variation methods for three-dimensional electrical resistance tomography inverse problems, *Water Resources Research*, *36*(7), 1653–1664.

Zhdanov, M., and E. Tolstaya (2004), Minimum support nonlinear parametrization in the solution of a 3D magnetotelluric invese problem, *Inverse Problems*, *20*, 937–952.