

Edge Preserving Imaging

C. F. YOUZWISHEN AND M. D. SACCHI*

Department of Physics, University of Alberta,
Abadh Bhatia Physics Laboratory,
Edmonton, AB, T6G 2J1, Canada

Journal of Seismic Exploration, 2006, Volume 15, 4, 45-58,

*Email: sacchi@phys.ualberta.ca

Introduction

Ill-posed geophysical inverse problems require a regularization term to recover a unique, stable solution. Quadratic regularization terms result in linear solutions that are computationally easy to solve. However, these regularization terms apply homogeneous smoothing to the recovered image and, consequently, they will tend to blur sharp material properties boundaries.

Blocky images can be recovered by non-quadratic regularization terms that enforce simplicity or sparseness on the model gradient. In this paper, we study the problem of designing Edge Preserving Regularization (EPR) strategies for inverse problems that arise in the context of linearized seismic inversion (imaging). In particular, we use the proposed EPR algorithm to reconstruct acoustic velocity perturbations in a multi-source, multi-receiver seismic experiment. The assumption is that the departure of the true velocity profile from the background model is small and can be modeled by a blocky or piecewise constant perturbation.

Traditional methods for inverting geophysical data rely on quadratic regularization techniques to find a unique and stable solution to an ill-posed problem (Tarantola, 1984; Titterton, 1985). When models have sharp material boundaries, quadratic regularization methods will tend to smooth the discontinuities. Edge preserving imaging seeks to reconstruct piecewise continuous profiles by enforcing a simplicity or sparseness constraint on the model gradient or on one or more spatial derivatives of the model parameters. The problem of finding regularization methods capable of preserving edges while suppressing artifacts due to noise has been studied in astronomy (Geman and Yang, 1995; Molina et al., 2001), medical imaging (Charbonnier et al., 1997; Barone, 1999), and geophysics (Last and Kubic, 1983; Farquharson and Oldenburg, 1998; Portniaguine and Zhdanov, 1999).

Edge preserving imaging in digital image processing and medical tomography has been implemented by describing edges via Markov random fields (MRF's) where model smoothing

is limited to operate on homogeneous areas, and is switched off in areas with edges (Geman and Geman, 1984; Chalmond, 1988; Geman and Reynolds, 1992).

In this paper we concentrate on the application of edge preserving regularization to the problem of estimating acoustic velocity perturbations from a multi-source, multi-receiver geophysical experiment. The problem is linearized using the single scattering Born approximation about a known reference medium (Beylkin, 1985; Miller et. al, 1987). We will assume that the departure of the true velocity profile from the background model is small and that it can be modeled by blocky perturbations. The results of the EPR algorithm are compared with images recovered using zero order quadratic regularization.

Inversion with Edge Preserving Regularization

Reconstruction of 2D acoustic profiles

In this paper we are concerned with the solution of the linear inverse problem that arises in the multi-source multi-receiver seismic acoustic probe in the Born approximation (Taylor, 1972):

$$\mathbf{u}_{sc} = \mathbf{L}\mathbf{f} + \mathbf{n} \tag{1}$$

where \mathbf{u}_{sc} indicates the measured scattered wavefield at the surface of the earth, \mathbf{f} denotes the acoustic potential and \mathbf{n} additive uncorrelated noise.

The classical regularization for geophysical inversion schemes is a quadratic regularization term or Tikhonov regularization (Tikhonov and Gorcharsky, 1985). In all our examples, we will compare the edge preserving regularization scheme to Tikhonov or quadratic regularization.

The cost function for the ill-posed inverse problem is written in terms of the acoustic potential, \mathbf{f} , as

$$J(\mathbf{f}) = J_D(\mathbf{f}) + \mu J_R(\mathbf{f}). \quad (2)$$

The first term represents the data misfit portion of the cost function

$$J_D(f) = \|\mathbf{L}\mathbf{f} - \mathbf{u}_{sc}\|_2^2, \quad (3)$$

while the second term, $J_R(\mathbf{f})$, is the regularization term.

In quadratic regularization, stability of the solution is attained by introducing a regularization term of the following form:

$$J_R = \mathbf{f}^T \mathbf{f} \quad (4)$$

The hyper-parameter, μ , indicates the priority of satisfying the regularization term over the data misfit term. In this case, we are not only inverting a model that fits the data to a certain tolerance but also that has minimum norm. The above regularization term is often referred as zero order regularization. Bear in mind that other quadratic regularization methods can be adopted by using a weighted norm of the form

$$J_R = \mathbf{f}^T \mathbf{W}^T \mathbf{W} \mathbf{f}. \quad (5)$$

In general, \mathbf{W} can be taken as a matrix of first order discrete derivatives (first order quadratic regularization) or a second order discrete derivative (second order quadratic regularization) (see, for instance, Constable et. al, 1987).

Edge preserving regularization

In edge preserving regularization the model norm takes the following functional form (Last and Kubic, 1983; Geman and Geman, 1984; Portiaguine and Zhdanov, 1999)

$$J_{EP}(\mathbf{x}) = \sum_i x_i^2 / (1 + x_i^2) \quad (6)$$

where \mathbf{x} is an arbitrary vector. The above function is also a measure of sparseness or simplicity. In other words, it can be used to produce models that are sparse. It is clear that in the present formulation of the problem, we seek a blocky solution not a sparse one. This is solved by considering the non-quadratic norm J_{EP} applied to the spatial derivatives of the model parameters. By forcing the spatial derivatives to be sparse we are automatically forcing the model parameters to be piecewise constant (blocky).

In our numerical experiments we have adopted the following norm

$$J_R(\mathbf{f}) = \sum_{i=1}^{nx} \sum_{j=1}^{nz} \mu_x J_{EP}([\mathbf{D}_x \mathbf{f}]_{i,j}) + \mu_z J_{EP}([\mathbf{D}_z \mathbf{f}]_{i,j}). \quad (7)$$

Here, \mathbf{f} is the discrete 2D acoustic potential discretized in a 2-D array of size nx by nz . The elements of first order finite difference derivatives are defined as

$$\begin{aligned} [\mathbf{D}_x(\mathbf{f})]_{i,j} &= (f_{i+1,j} - f_{i,j})/\delta \\ [\mathbf{D}_z(\mathbf{f})]_{i,j} &= (f_{i,j+1} - f_{i,j})/\delta. \end{aligned} \quad (8)$$

The constant cell width, Δ , traditionally placed in the derivative terms is absorbed into the scaling hyper-parameter, δ . To reduce the number of indexes, the 2D model derivatives are expressed in lexicographic notation as

$$J_R(\mathbf{f}) = \sum_{l=1}^{n_x \cdot n_z} \mu_x J_{EP}([\mathbf{D}_x \mathbf{f}]_l) + \mu_z J_{EP}([\mathbf{D}_z \mathbf{f}]_l). \quad (9)$$

The minimum of the cost function of the problem (equation (2)) with J_R given by expression (9) is found via the following iterative scheme (Scales, 1987)

$$\tilde{\mathbf{f}}^k = \left[\mathbf{L}^T \mathbf{L} + \mu_x \mathbf{D}_x^T \mathbf{Q}_x^k \mathbf{D}_x + \mu_z \mathbf{D}_z^T \mathbf{Q}_z^k \mathbf{D}_z \right]^{-1} \mathbf{L}^T \mathbf{u}, \quad (10)$$

with the elements of the diagonal weighting matrices given by:

$$[\mathbf{Q}_x^k]_{ll} = \frac{1}{(1 + [\mathbf{D}_x \mathbf{f}^{k-1}]_l^2)^2}, \quad (11)$$

and

$$[\mathbf{Q}_z^k]_{ll} = \frac{1}{(1 + [\mathbf{D}_z \mathbf{f}^{k-1}]_l^2)^2}. \quad (12)$$

The solution $\tilde{\mathbf{f}}^k$ is the recovered solution at iteration k . The convergence of this type of algorithms is analyzed by Charbonnier et al. (1997); and convergence from any initialization to a local minimum is guaranteed. We run the previous algorithm until $\|\tilde{\mathbf{f}}^{(k+1)} - \tilde{\mathbf{f}}^k\|_2^2 / \|\tilde{\mathbf{f}}^k\|_2^2 < \epsilon$ where ϵ is a small constant. It is clear that the process must be repeated for different values of the hyper-parameters (μ, δ) until a solution with desirable features is obtained. We will come back to this point when describing our numerical simulations.

The diagonal weighting matrices, \mathbf{Q}_x and \mathbf{Q}_z are responsible for locating the presence of edges, and adjusting the amount of smoothing applied to the solution. By examining equations (11) and (12), it is easily seen that if the derivative of the model is small (no edge), the diagonal weighting term will approach 1, and full smoothing is applied. Conversely, if

the derivative of the model is large (an edge is present), the diagonal weighting term will approach zero, and the smoothing term is turned off. In this way, the algorithm will mark the location of the edges, and adjust the amount of smoothing in order to preserve them.

The discontinuities are preserved according to the threshold value chosen to specify an edge as opposed to a small artifact. This threshold value is controlled by the scaling parameter δ (see Figure 1). In the background areas of the model, the algorithm is able to apply homogeneous smoothing and encourage a piecewise smooth solution.

Examples

To explore the behavior of the two hyperparameters, EPR will first be tested on a 1-D model. Assuming a zero-offset experiment, or one where the source and receiver are in the same location, the seismic waveform data of a horizontally homogeneous Earth are measured. The velocity model, source wavelet, and synthetic data are displayed in Figure 2. The background velocity is chosen to be a homogeneous 2000 m/s, and the acoustic perturbation is then computed to create synthetic data. The initial model is set to a homogeneous vector of zeros.

This simple experiment reveals that multiple pairs of the parameters μ and δ will give the same value of χ^2 , our test for goodness of fit (Figure 3 A). The second part of the figure illustrates three solutions sharing the same χ^2 value. Only parameters in a limited range will yield a solution that is piecewise constant.

Finally, Figure 3 C displays the effect that the scaling parameter, δ , has on the solution when μ is held constant. A large scaling parameter will result in small values for the scaled derivative. The weighting term will approach one, and the smoothing operator receives near full weight. Therefore, the solution of a large δ value will approach a smoothed solution. The Tikhonov regularization of the acoustic potential is included for comparison (with J_R given by

equation (4)). On the other hand, when the scaling parameter is small, the scaled derivative will be magnified. The weighting term will approach zero to turn off smoothing, and preserve the discontinuities. A very edgy solution is expected for small scaling parameters.

Discontinuities in the solution are not controlled solely by the scaling parameter. They are also influenced by the weighting parameter, μ . If the initial model is set to be homogeneous, no edges are detected in the first iteration of the algorithm. Therefore, full smoothing is applied to the model, and the weighting parameter for this iteration becomes a combination of the two parameters

$$\mu_z^1 = \frac{\mu_z}{\delta}, \quad (13)$$

where μ_z^1 denotes the vertical weighting parameter for the first iteration of the algorithm. It is apparent that a large value of μ_z , or a small value of δ will create a large weighting parameter that will enforce a very smooth solution. It is this combination of parameters that creates the initial solution that the edges are detected and classified from. Figure 4 A shows the progression of both the solution and the weighting matrix as the algorithm progresses for the edgy solution ($\delta = 0.01$) obtained in Figure 3 C. The small value of the scaling parameter combines with the weighting parameter to enforce a very smooth solution in the first iteration. Thus, when the edges are identified and preserved, there are not as many as one would expect. The progression of the algorithm for the optimum solution ($\delta = 1$) and the smooth solution ($\delta = 0.1$) of the same problem are shown in Figures 4 B and 4 C.

To further test the algorithm, a more complicated 2D velocity model is tested (Figure 5 A). Four sources at 50 m spacing are placed at the surface, and 20 receivers are spaced every 10 m at the surface. The background velocity model is set to a homogeneous 2300 m/s. Again, the initial model is set to a homogeneous vector of zeros, and a total of 5 percent Gaussian noise is added to the synthetic data.

The solutions retrieved by Tikhonov inversion and EPR are displayed in Figure 5 C and D. The EPR inversion converges in 5 iterations of the algorithm. Figure 6 shows the progression of the updated model as it is refined by the algorithm. The weighting matrices, \mathbf{Q}_x and \mathbf{Q}_z , clearly mark the positions of the edges, and turn off the smoothing at these locations. The EPR technique recovers an almost perfect solution, the exception being made at the horizontal edges of the model. At these locations there is not enough data coverage to resolve the model, due to the source-receiver geometry.

Thus far, only homogeneous background models have been considered. The algorithm is easily adapted for non-homogeneous background models: all that are necessary are traveltimes and amplitude tables in order to calculate the background Green's functions. A four layer background velocity model is created, and the acoustic perturbation used in the previous example is superimposed on top (Miller et. al, 1987). The final inverted acoustic perturbation is added to the background velocity model to retrieve the final solution.

The velocity model, synthetic data, and its solutions are displayed in Figure 7. The EPR solution converges within 5 iterations of the algorithm. Again, the superiority of the algorithm for recovering high resolution images is shown.

Conclusions

Standard quadratic regularization is not appropriate for recovering blocky images. Quadratic Tikhonov regularization constrains the solution around a reference value, and applies homogeneous smoothing to the complete image. This means one must compromise between edge-preservation and suppression of highly oscillatory features arising from the inversion of noisy data.

In contrast, non-quadratic edge-preserving regularization can apply a large degree of smoothing without smearing the boundaries of the solution. The weighting matrices of the problem

detect the presence of edges and preserve them by turning off smoothing at the appropriate locations. This is like introducing an adaptive trade-off parameter. Thus, there is no compromise between the amount of smoothing applied, and the edge-preservation.

EPR requires more effort to implement than Tikhonov regularization because it has two hyperparameters rather than one. As well, EPR is computationally more expensive. The solution specified by Tikhonov regularization is recovered by solving one linear inverse problem. In contrast, EPR requires 5-6 iterations of a weighted, linear solution. Because of these considerations, EP regularization could be applicable to small scale geophysical studies, such as vertical seismic profiles (VSP's), or near surface environmental seismic imaging problems.

Acknowledgments

This research has been supported by grants from NSERC-Canada, the Alberta Department of Energy, and the following industrial sponsors: Geo-X Ltd., EnCana Petroleum Ltd., and Veritas Geoservices.

References

- Barone, P., 1999, Fast deconvolution by a two-step method: *SIAM J. Sci. Comput.*, **21**, 883-899.
- Beylkin, G., 1985, Imaging of discontinuities in the inverse scattering problem by inversion of a casual generalized Radon transform: *J. Math Phys.*, **26**, 99-108.
- Chalmond, B., 1988, Image Restoration using an estimated Markov model: *Signal Processing*, **15**, 115-129.
- Charbonnier, P., Blanc-Feraud, L., Aubert, G., and Barlaud, M., 1997, Deterministic Edge-Preserving regularization in computer imaging: *IEEE Trans. Image Proc.*, **6**, 298-311.
- Constable, S. C., Parker, R. L., and Constable, C. G., 1987, Occam's Inversion: A practical algorithm for generating smooth models from electromagnetic sounding data: *Geophysics*, **52**, 289-300.
- Farquharson, C. G., and Oldenburg, D. W., 1998, Non-linear inversion using general measures of data misfit and model structure: *Geophys. J. Int.*, **134**, 213-227.
- Geman, S., and Geman, D., 1984, Stochastic relaxation, Gibbs distributions, and Bayesian restoration of images: *IEEE Trans. Pattern Anal. Machine Intell.*, **6**, 721-741.
- Geman, D., and Reynolds, G., 1992, Constrained restoration and the recovery of discontinuities: *IEEE Trans. Pattern Anal. Machine Intell.*, **14** 367-383.
- Geman, D., and Yang, C., 1995, Nonlinear image recovery of half-quadratic regularization: *IEEE Trans. Image Proc.*, **4**, 932-945.
- Li, Q., 2001, High Resolution Hyperbolic Radon Transform Multiple Removal: MSc Thesis, Department of Physics, University of Alberta, Edmonton, Canada.

- Miller, D., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: *Geophysics*, **52**, 943-964.
- Molina, R., Nunez, J., Cortijo, F. J., and Mateos, J., 2001, Image Restoration in Astronomy: A Bayesian Perspective: *IEEE Signal Proc. Mag.*, **18**, 11-29.
- Portniaguine, O., and Zhdanov, M. S., 1999, Focusing geophysical inversion images: *Geophysics*, **64**, 874-887.
- Sacchi, M. D., and Ulrych, T. J., 1995, High-resolution velocity gathers and offset space reconstruction: *Geophysics*, **60**, 1169-1177.
- Sacchi, M. D., 1997, Re-weighting strategies in seismic deconvolution: *Geophys. J. Int.*, **129**, 651-656.
- Scales, J. A., 1987, Tomographic inversion via the conjugate gradient method, *Geophysics*, **52**, 179-185.
- Tarantola. A., 1984, Inversion of seismic reflection data in the acoustic approximation: *Geophysics*, **49**, 1259-1266.
- Taylor, J. R., 1972, *Scattering Theory*: Wiley, New York.
- Tikhonov, A. H., and Gorcharsky, A. V., 1987, *Ill-posed problems in the natural sciences*: Mir Publications.
- Titterigton, D. M., 1985, General structure of regularization procedures in image reconstruction problems: *Astron. Astrophs.*, **144**, 381-387.

List of Captions

1. A 1D depth model, the scaled derivative, and weight on the smoothing term for three different values of the scaling parameter, δ . As δ becomes small, the derivatives are magnified, and all discontinuities are classified as edges. Smoothing at these locations is minimized, indicated by the near-zero weights. As the scaling parameter becomes large, none of the discontinuities qualify as edges, and the weights are near 1 indicating full smoothing.
2. The (A) velocity model, (B) source wavelet, and (C) synthetic data.
3. (A) A logarithmic plot of χ^2 versus μ for different values of δ . (B) Solutions for different parameter pairs of δ , and μ , all sharing the optimum χ^2 value for this problem. (C) The initial model, solutions for 3 values of δ , and the Tikhonov (TK) regularization, where $\mu = 10^{-10}$.
4. The solution, and values of the diagonal weighting matrix (dashed line) as the EPR algorithm progresses from the second to the fifth iteration of the (A) edge, (B) optimum, and (C) smooth solutions. The weighting terms have a maximum value of one for full smoothing, and a minimum of zero to turn off smoothing. Note that the first iteration is not shown because the solution is set to zero, and the weighting terms are all unity to enforce full smoothing.
5. The (A) velocity model, (B) synthetic data, the solution using (C) Tikhonov regularization, and (D) EPR solution. The parameters corresponding to these solutions are: (C) $\mu = 200$, and (D) $\delta_x = \delta_z = 0.35$; $\mu_x = \mu_z = 8$.
6. The diagonal weighting matrices \mathbf{Q}_x and \mathbf{Q}_z , and the updated model at iterations 1, 3, and 5 of the EPR algorithm. At the first iteration, the model is homogeneous, so the weighting matrices apply full smoothing by default.
7. The (A) layered velocity model, (B) synthetic data, and solutions using (C) Tikhonov

regularization and (D) EPR solution. The parameters corresponding to these solutions are: (C) $\mu = 100$, and (D) $\delta_x = \delta_z = 0.5$; $\mu_x = \mu_z = 50$.

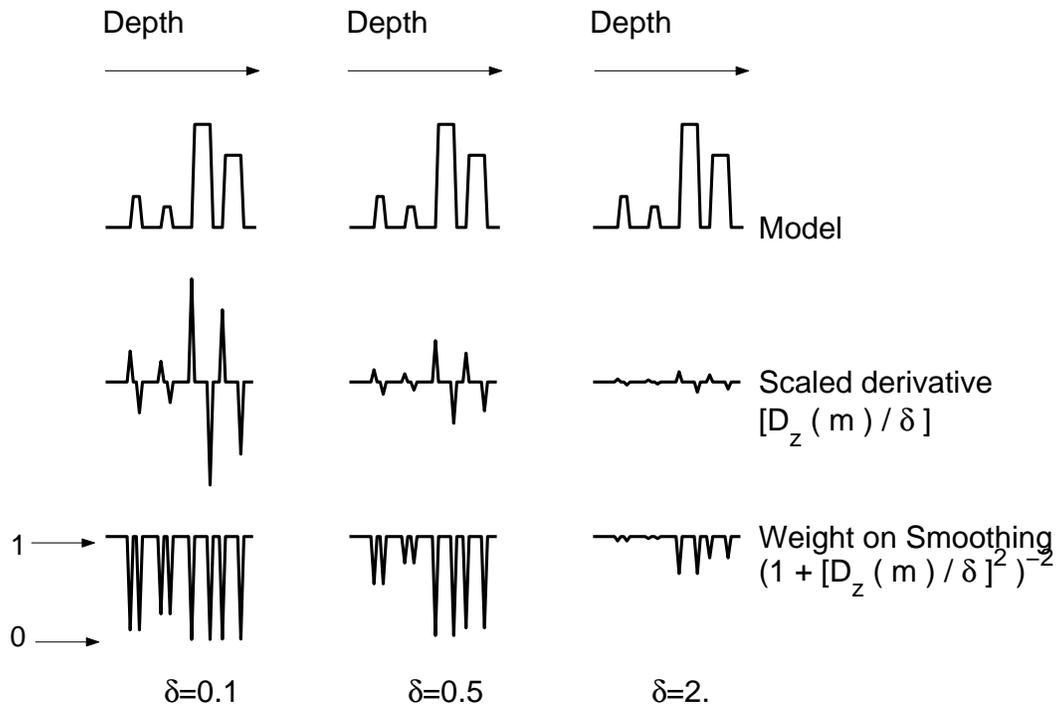


Figure 1:

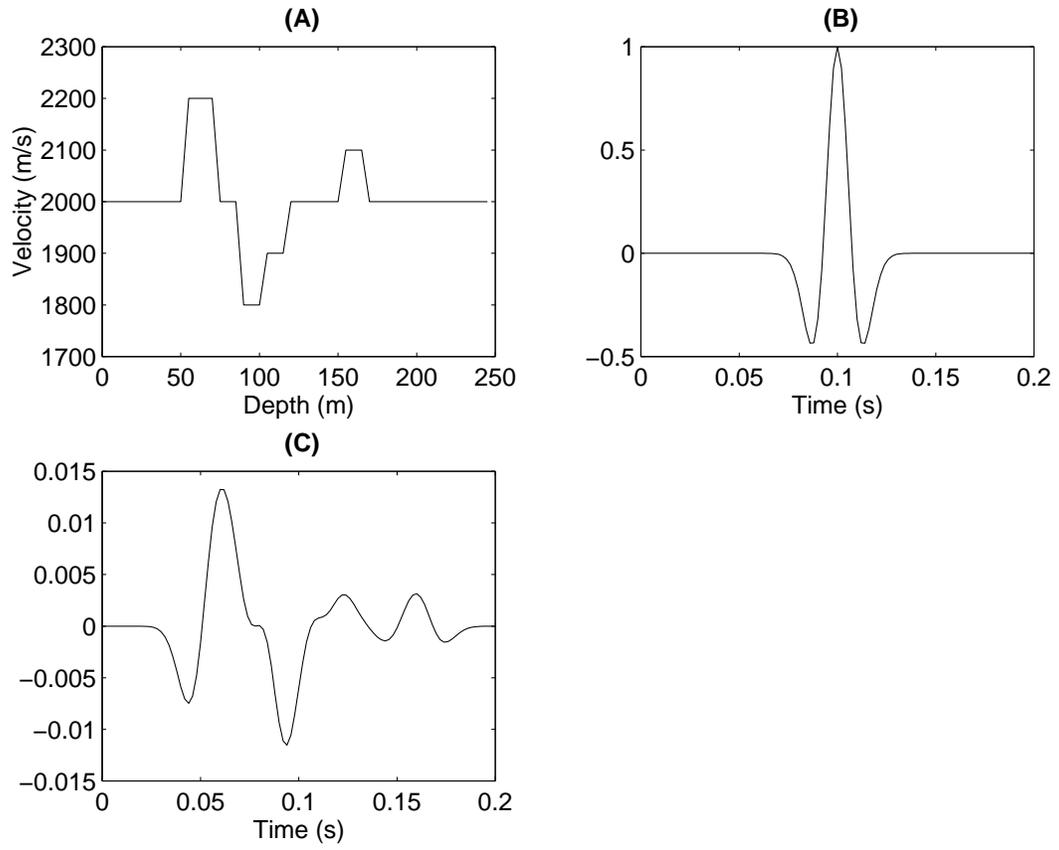


Figure 2:

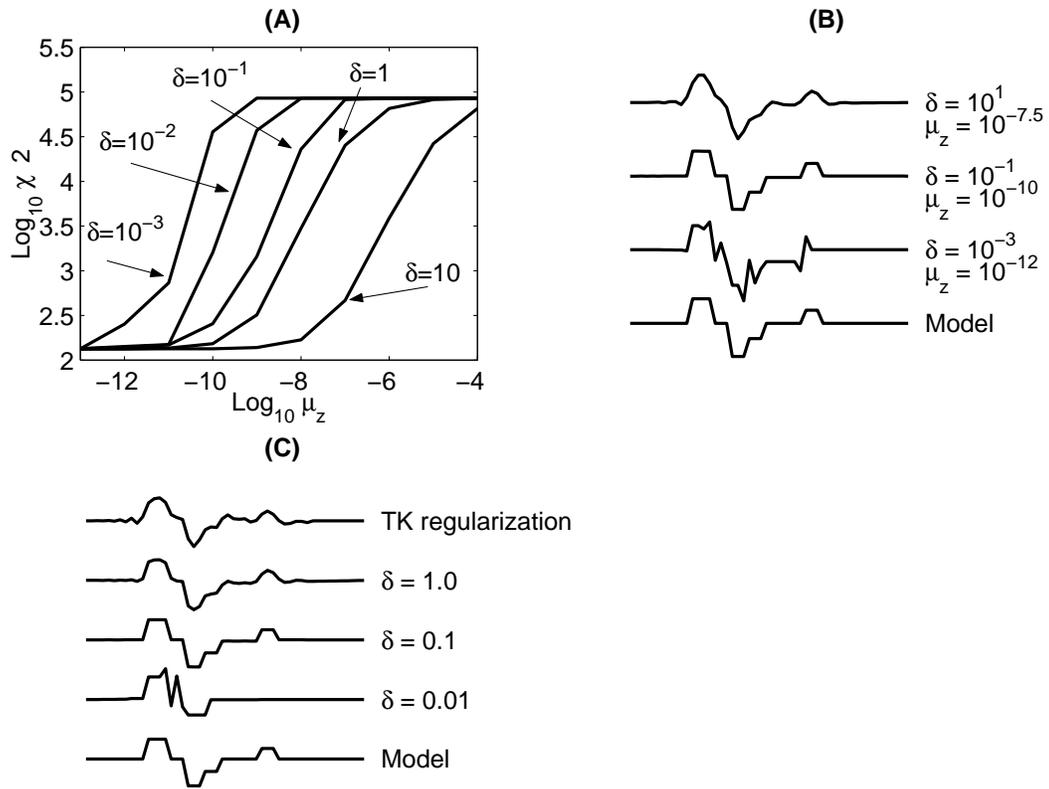


Figure 3:

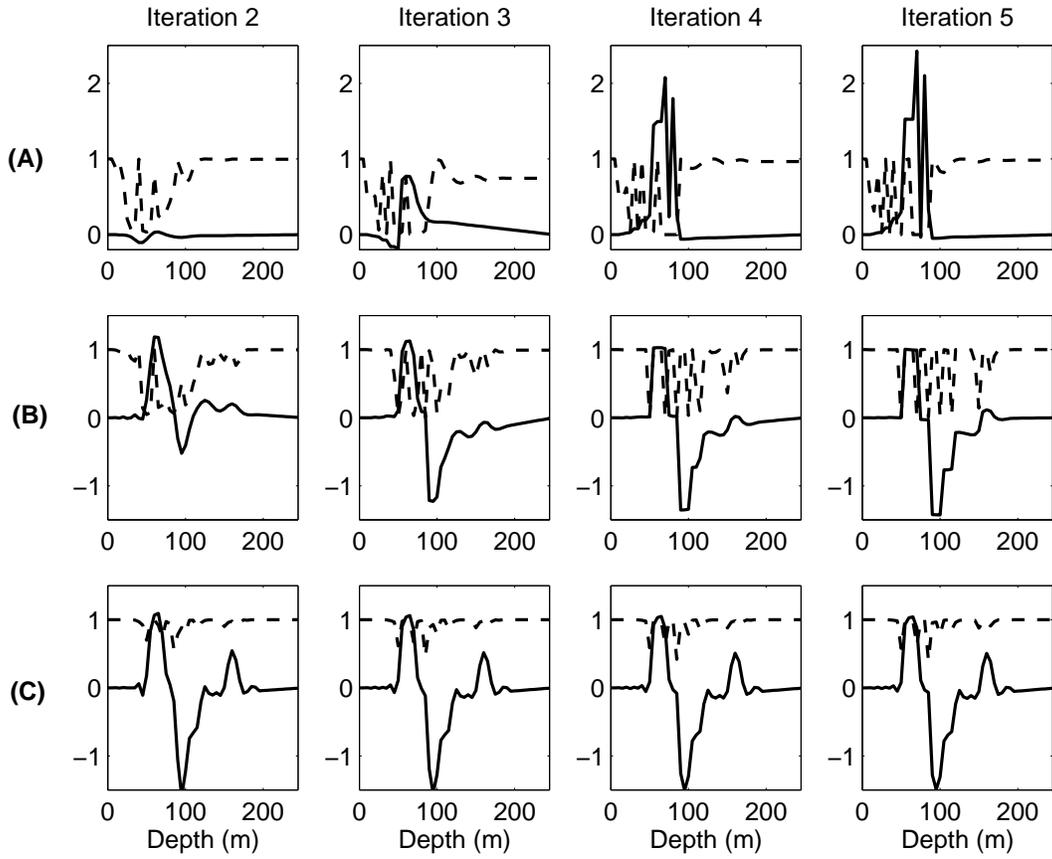


Figure 4:

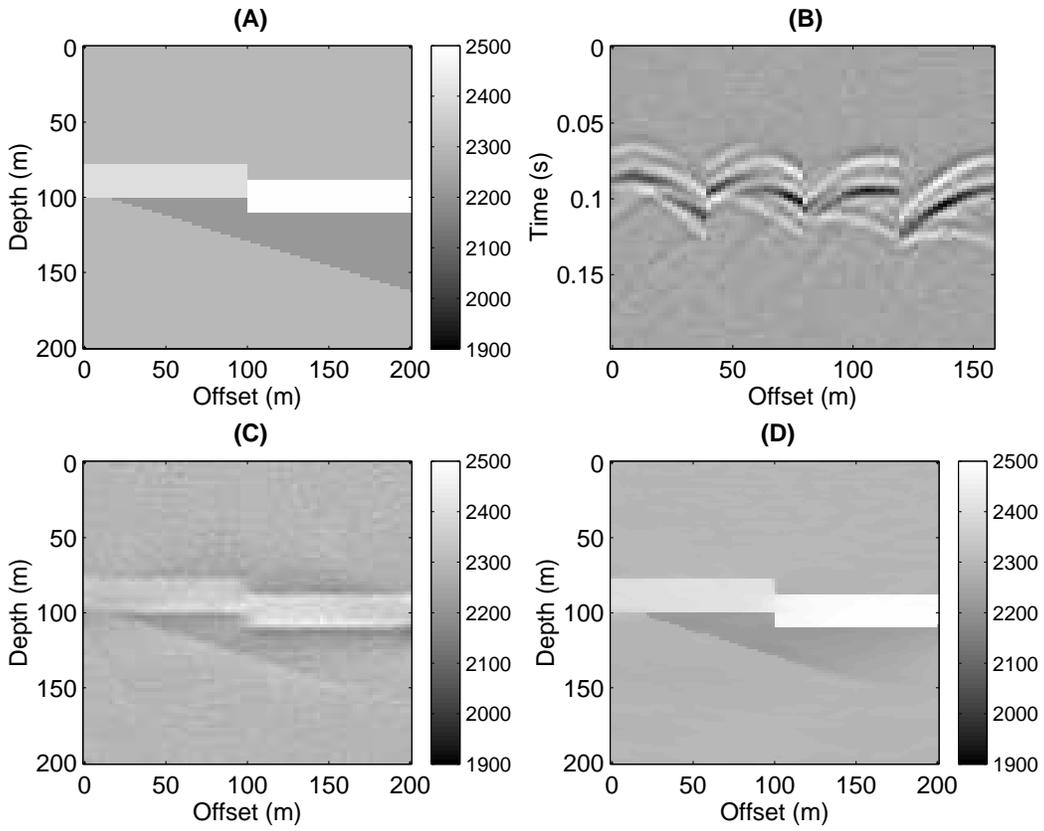


Figure 5:

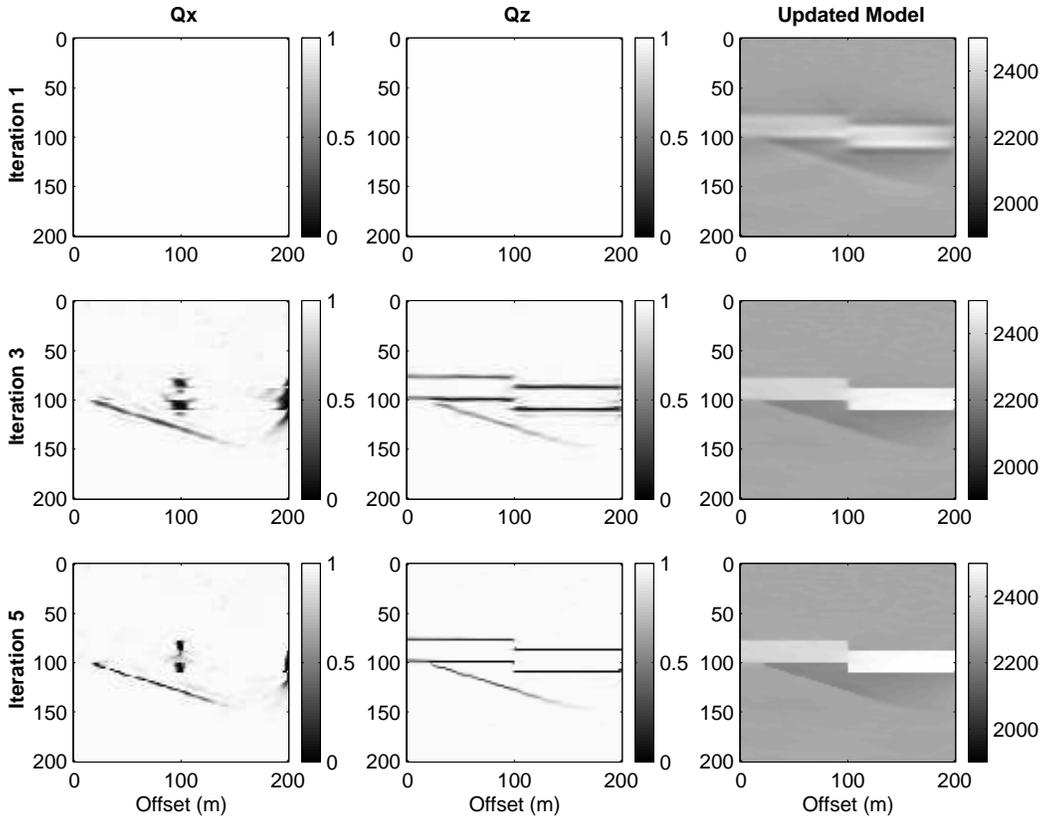


Figure 6:

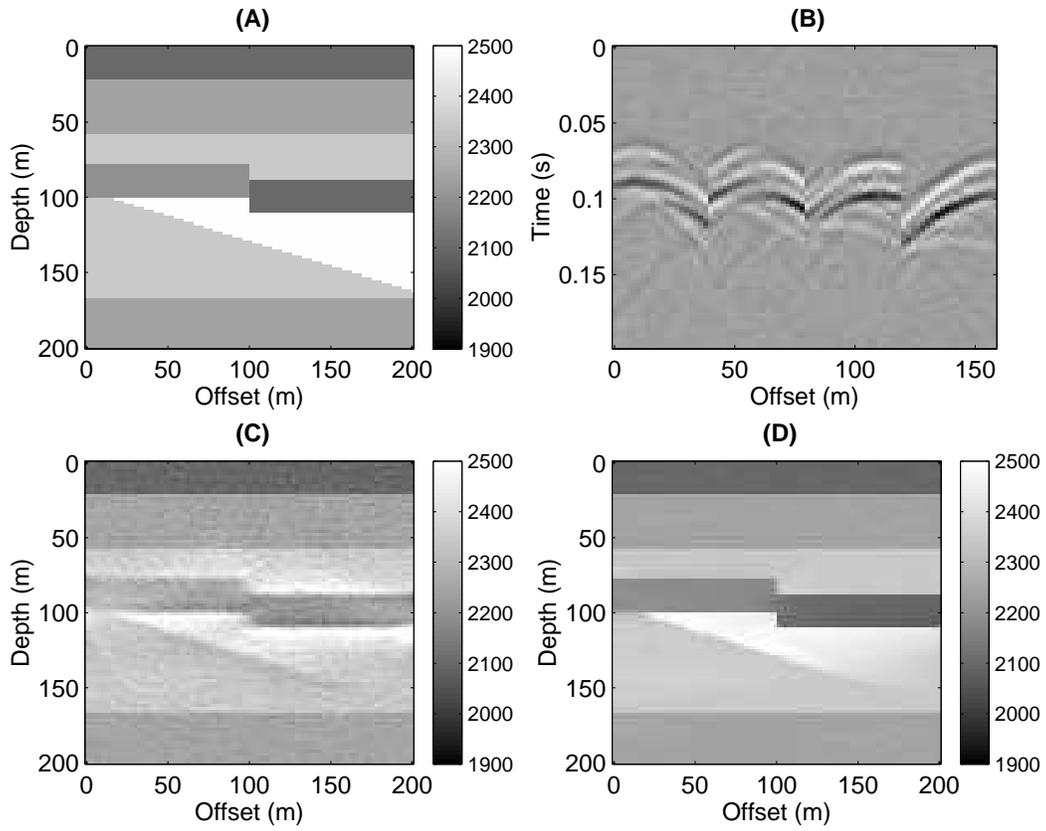


Figure 7: