NONLINEAR WAVEFORM INVERSION OF SEISMIC REFLECTION DATA: THE MARMOUSI MODEL

Mark Noble and Albert Tarantola

*Institut de Physique du Globe de Paris
4, Place Jussieu
75252 Paris Cedex 05
France*

ABSTRACT

Data from 45 shot records from the *Marmousi* data set collected in a conventional, marine seismic reflection survey are inverted to obtain the subsurface map of the short wavelength variations of the impedances. The inversion method employs a nonlinear, iterative, gradient algorithm which inverts waveform data; so there is no picking of travel times or amplitudes. The algorithm is computationally expensive as it relies heavily upon elastic finite difference modeling. In addition, because a gradient search is employed to minimize the misfit between observed and modeled data, we attempt to solve only the weakly nonlinear problem associated with the short wavelengths of impedance (information on the short wavelengths of impedance is mainly derived from amplitude variations in the data). Resolving the long wavelengths of velocity (background) is thought to be a highly nonlinear problem, at least for waveform inversion, and is not attempted herein. For the inversion of this data set, long wavelengths have been derived using the *focusing analysis* optimization technique developed by Total-CFP.

INTRODUCTION

In a simple, straightforward, and ideal formulation of the seismic inverse problem one finds the set of model parameters which best predict synthetic seismograms that *best fit* the observed seismograms. The set of model parameters should include all unknowns of the problem such as elastic stiffness,

density, attenuation coefficients, source functions, receiver responses, etc. This philosophy is highly dependent on our ability to accurately model wave propagation and is not equivalent to the philosophy of traditional seismic data processing. In the formulation of nonlinear waveform inversion the solution is derived iteratively such that we model synthetic seismograms at each iteration and thus have a quantitative measure of the reliability of our results. We are also able to qualitatively identify events in our observed data and associate them with our resulting subsurface models. If the inversion does not resolve certain events, we know which parts of our model to distrust. Such a qualitative analysis can be an important tool for the seismic interpreter. Unfortunately, we do not, at present, have sufficient computer power to statistically quantify our results. This is simply because the problem is so large. For the particular results we present here, there are over 125,000 unknowns and approximately 1.2 million observed data values.

In any case, it is mainly the advent of the supercomputer which allows us to step away from traditional data processing and move toward solving the inverse problem; however, we are still limited in our quest for a complete solution of the inverse problem, at least for the next few years, because, in addition to our inability to statistically quantify our results, we cannot yet economically model all aspects of wave propagation. Although the world of seismic modeling is progressing rapidly, we have no

choice, at present, but to approximate those things that we cannot properly model. We will discuss our various approximations as we proceed. For the results we present herein, we model wave propagation using high-order, in both space and time, finite difference solution to the two-dimensional heterogeneous, wave equation. For a detailed discussion of this modeling scheme, the interested reader is referred to Crase et al. [1990a,b]. Obviously, for a complete solution of the inverse problem one should model three-dimensional wave propagation which also includes the effects of attenuation and anisotropy. However, even in our two-dimensional formulation, computer resources prevent us from inverting for the highly variable impedance distribution of the near surface.

Conventional seismic exploration data contains two distinct types of information [Claerbout, 1985; Jannane et al., 1990]. Information about the long wavelengths (typically greater than 200 m) of propagation velocity is associated with travel time, while information concerning the short wavelengths (typically less than 80 m) of impedance is associated with reflection amplitude. There appears to be a null space between the two types of information such that conventional data are insensitive to impedance and velocity variations with wavelengths between 80 and 200 m (the exact range, of course, depends on data acquisition parameters).

These results, as simple as they may appear, are important. They are more or less known by the seismic community but, to our knowledge, no quantitative discussion has previously been published. The fact that long and short resolvable wavelengths are separated by a region of non resolvable wavelengths, suggests that we can parameterize differently the long wavelengths of velocities and the short wavelengths of impedances. For instance, a few tens of nodes will suffice to define the velocities, while a very dense grid of points is necessary to define the image of impedance contrasts. The parameters describing the velocity are strongly nonlinearly related to the data, while the parameters describing the impedance contrasts are almost linearly related to the data.

Our research has shown that gradient methods of optimization work perfectly for retrieving the image of impedance contrasts, but do not perform as well for the velocity background. Ideally we should use a fully nonlinear method to optimize for the velocity background such as Monte Carlo or Simulated Annealing (Koren et al. [1990], Mosegaard and Tarantola [1989]), and gradient methods to optimize for the impedance contrasts. This is very similar to the two-step industrial method of velocity analysis plus migration.

DATA SPECIFICATIONS

For the 45 shot records used in the inversion, the data acquisition consisted of 60 hydrophone arrays. The near-trace and far-trace offsets are 200 and 1675 m, respectively, with 25 m intervals between hydrophone groups. Each array covered 25 m which we simulate in our finite difference modeling by summing over appropriate grid points. The source was a "water gun" and the source arrays covered 40 m (also simulated in the modeling). The shot point interval is 50 m (shot point 90 to shot point 180). Source depth is 8 m and receiver depth is 10 m. The recording sample interval was 4 ms, resampled to 1 ms to accommodate the finite difference modeling, and we used 1.5 s of data for the inversion. Figure 1 shows an example of the 45 shot records after scaling by $t^{1.5}$ and muting the first arrivals. This record was recorded at shot point 146, and is quite typical of all shot records of the data set we are inverting, that is the shooting was done over a complex 2 dimensional structure.

DATA PREPROCESSING

Because we do not fully understand the nonlinear aspects of head waves with respect to waveform inversion, we applied a first arrival mute to the data. During the inversion we also mute our modeled data with the same mute function.

When inverting real data, we would apply several more preprocessing steps before it was input to the inversion algorithm. First, because our inversion algorithm is 2D and wave propagation for the real data is in a 3D medium, we would scale the observed data by $\sqrt{t}$ to approximate a change from spherical to cylindrical spreading. Second, be-

cause of our inability to economically model the fine detail of the near surface, we would apply surface-consistent static corrections in an attempt to lessen the effects (at least partially) caused by the high variability of near-surface velocities. This step would only be performed for land data. Our final preprocessing step would be to calculate and apply scalars to each trace in an attempt to at least crudely correct for inconsistent geophone and source coupling, which can sometimes be observed.

THE INITIAL MODEL

In our work, we have used an iterative gradient algorithm to solve the weakly nonlinear problem associated with the short wavelengths of the impedance. We do not invert for the long wavelengths of the background velocity. The works of Tarantola [1986] and Landa et al. [1989], and also our own experience, indicate that retrieving the background velocity via waveform inversion can be a highly nonlinear problem and thus may not be adaptable to the gradient methods which we employ. We also note that we do not invert for the density because of its difficult resolvability in the absence of data collected at very long offsets [e.g., see Beydoun and Mendes, 1988]. A complete solution of the problem would not only include density but would also simultaneously invert for the background velocities. We feel that the solution for the background velocities requires a "Monte-Carlo" type search of the model space, and the interested reader is referred to the work of Cao et al. [1990] who present a linearized example of a simultaneous approach to the problem. Although such a theory can be readily formulated for nonlinear solution using a simultaneous approach, it is again beyond our present computing capabilities.

With regard to initial model of subsurface parameters, the short wavelengths of P impedance were set to zero. The P velocity background model of our inversion was obtained in collaboration with Total-CFP using a *focusing analysis* optimization technique. Figure 2 shows one example of the velocity profiles obtained and is very close to the true model: figure 3. To construct a density background, we used an empirical relationship with P velocity (Gardner et al., 1974) in addition to the density logs

given with the data set. Once we had derived the P velocity and density background, they were converted to impedance and density for the actual inversion. We again point out that the background model was not allowed to change during the inversion.

We consider our estimation of the source functions as part of constructing the initial model even though we did not attempt to retrieve source functions as part of this particular inversion. Although the theory of waveform inversion predicts that one should be able to recover the source, we do not fully understand at present whether or not inversion for the source is a sufficiently constrained problem. In any event we chose to estimate the source function before inverting the data and then to hold the source constant throughout the inversion.

As an estimate of the source, we used a Ricker wavelet with a central frequency of 25 Hz. As an estimate of the amplitude of the source, we modeled a single shot record, using the well-logs (velocity and density at position 9004) and derived a source scalar by comparing the average absolute amplitude of the synthetic record after first-arrival muting to the average absolute amplitude of the observed data.

Obviously, our means of estimating the source function are extremely crude and much more work on the topic of source functions is necessary.

THE INVERSION ALGORITHM

Before discussing our results, we will give a very brief outline of the inversion algorithm which we use for our nonlinear inversion. For a more detailed account of this algorithm, we refer the interested reader to Crase et al. [1990a,b].

In essence, the misfit function that we want to minimize is given by

$$S_{L_2}(\mathbf{m}_n) = \frac{1}{2} \sum_i \left(\frac{d^i - f^i(\mathbf{m}_n)}{\sigma^i} \right)^2, \quad (1)$$

where $\mathbf{m}_n$ are the model parameters at the n^{th} iteration, $\mathbf{d}$ is the observed data, and $f(\mathbf{m}_n)$ is synthetic data generated using the current model parameters. In practice, σ^i is some estimate of the error deviation corresponding to a given ($d^i -$

$f^i(\mathbf{m}_n)$). We note that for the *Marmousi* data, we have used the root-mean-square value of all samples of a particular shot record as the estimated error deviation for all samples of that shot.

To minimize the misfit, we use a preconditioned gradient subspace method (e.g., see Kennett and Williamson, 1987). In this method, a model update at the n_{th} iteration is given by

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \sum_j \alpha_n^j \phi_n^j, \quad (2)$$

$$\phi_n = -\mathbf{P}_n \hat{\gamma}_n,$$

where $\hat{\gamma}_n$ is the gradient of the misfit function with respect to all model parameters, $\mathbf{P}_n$ is the preconditioning operator, and α_n^j is the optimal steplength for the model change ϕ_n^j of the j^{th} subspace. As subspaces, we can choose P impedance, S impedance, the magnitude of each source array. The *Marmousi* data being acoustic and the source constant for all the shots, we had only one subspace: the P impedance. As a preconditioning operator, we apply a scaling function (with a spatially varying amplitude) to the gradient which attempts to adjust it for the energy disparities caused by recording geometry and wavefield spreading.

The gradient of the misfit function is

$$\hat{\gamma}_n = \mathbf{F}_n^T \hat{\mathbf{r}}_n, \quad (3)$$

where

$$\mathbf{F}_n = \frac{\partial f(\mathbf{m}_n)}{\partial \mathbf{m}_n}, \quad (4)$$

and $\hat{\mathbf{r}}^i$ are the weighted residuals given by

$$\hat{r}_n^i = \frac{1}{\sigma^i} \left(\frac{d^i - f^i(\mathbf{m}_n)}{\sigma^i} \right). \quad (5)$$

To determine the steplengths of each subspace, we assume linearity in the immediate vicinity of our current model parameters and solve a new minimization problem. Defining

$$a_n^i = d^i - f^i(\mathbf{m}_n), \quad (6)$$

and

$$b_n^{ij} = f^i(\mathbf{m}_n + \phi_n^j) - f^i(\mathbf{m}_n), \quad (7)$$

the new function to minimize with respect to the steplengths is given by

$$\Psi(\alpha_n^1, \dots) = \frac{1}{2} \sum_i \left(\frac{1}{\sigma^i} \left(a_n^i - \sum_j \alpha_n^j b_n^{ij} \right) \right)^2 \quad (8)$$

The minimum of this new function is unique and can be found by using relaxation methods. We note that although the gradient calculation uses a linear operator and our steplengths assume linearity in the immediate vicinity of the current model parameters, we are *not* performing a linearized inversion. As we see in equation (2) for the preconditioned gradient subspace method, the gradient only gives us a direction in which to update the model parameters, and the steplengths tell us the size of the update. Our methods are *fully nonlinear* by the fact that we begin each successive iteration with a new reference model; thus our misfit function has no linear constraints.

THE INVERSION RESULTS

The subsurface grid of P impedance for which we inverted has 4 m spacing vertically and 5 horizontally. For the finite difference modeling of individual shots, grid size was 350 points horizontally and 368 points vertically. We used a fourth order spatial and fourth order temporal scheme which takes approximately 9.5 minutes of CPU time per shot on a Convex C1. One complete iteration of the inversion takes approximately 21 hours of computation. This is of course very costly and we only performed one iteration. But the world of computing is progressing rapidly. At the moment of concluding this paper, we have installed the inversion algorithm on the new Connection Machine just arrived at our Institute. The first tests are very encouraging. The same modelling on the CM2 (with 32 K processors) takes 14 s, one complete iteration should not take more 40 minutes. The high degree of parallelism and speed of computation available on the CM 2 make it an ideal computer for this algorithm. (All the computation times given for the CM 2 include IO.)

As previously mentioned, we did not invert for the near-surface impedances; thus the inversion results range from 200 to 1350 m in depth. The lateral

extent is limited to approximately 3000 m where we feel there is good multiplicity in subsurface coverage. The P impedance map after the first iteration is shown in figure 4. The initial background impedance has been removed. As pointed out by Tarantola [1988], one iteration of the algorithm is very similar to a reverse time, prestack migration. Note, however, that the subsurface image of the inversion is in terms of real parameters (impedance) unlike migration that has poorly defined units (reflectivity, etc.).

In figure 5, we show the modeled data, which corresponds to the same shot record (observed data) as in figure 1. Note that the synthetic data is modeled in the updated subsurface medium as defined in equation (2). All shot records (modeled, and observed) have been plotted at the same gain for easy comparison.

Obviously, after only one iteration, we cannot draw many conclusions or claim that we have converged to a solution. After the first iteration, we seem to have found the strong impedance contrasts and a few more iterations are certainly necessary to recover the weaker ones. A comparison of the modeled data to the observed data qualitatively shows which parts remain unexplained by the inversion algorithm. The main difference between the two data sets, is the ringing aspect of the modeled data which is due to the ghosts of the source and should disappear with a few iterations (from our own experience, 4 or 5 iterations are enough).

CONCLUSIONS

Although the problem of inverting *real* data involves many unknowns and data values (over 125,000 unknowns and approximately 1,2 million data values), nonlinear waveform inversion of conventional marine seismic data is a reality. It is mainly the advent of faster supercomputers which allows us to step away from traditional data processing and move towards solving the inverse problem. Obviously, waveform inversion is still expensive and is in its infancy. Much work lies ahead.

ACKNOWLEDGEMENTS

The support from the sponsors of the Geophysical Tomography Group (Amoco, Arco, Aramco,

CGG, Conoco, Convex, Digital, Elf-Aquitaine, Exxon, IFP, ITA, Mobil, Shell, Sun Microsystems and Total-CFP) is greatly acknowledged.

We kindly thank TOTAL-CFP for giving us a very good initial velocity model.

REFERENCES

- Beydoun, W. B., and Mendes, M., 1989, Elastic ray-born l2 migration/inversion, *Geophysics J.*
- Cao, D, Beydoun, W., Singh, S., and Tarantola, A., 1990, A simultaneous inversion for background velocity and impedance maps, *Geophysics*, **55**, 458-469.
- Claerbout, J. F., 1985, Imaging the Earth's interior, Blackwell Scientific Press.
- Crise, E., Pica, A., Noble, M., McDonald, J. and Tarantola, A., 1990a, Robust elastic nonlinear waveform inversion: application to real data. *Geophysics*, **55**, 458-469.
- Crise, E., Wideman, C., Noble, M., and Tarantola, A., 1990b, Nonlinear Elastic Waveform Inversion of Land Seismic Reflection Data, *J. Geophys. Res.*, (accepted, in press).
- Gardner, G. H. F., Gardner, L. W. and Gregory, A. R., Formation velocity and density—The diagnostic basis of stratigraphic traps, *Geophysics*, **39**, 770-780, 1974.
- Jannane, M., Beydoun, W., Crise, E., Cao, D., Koren, Z., Landa, E., Mendes, M., Pica, A., Noble, M., Roeth, G., Singh, S., Sneider, R., Tarantola, A., Trezeguet, D., and Zie, M., 1989, Wavelengths of earth structures that can be resolved from seismic reflection data, *Geophysics*, **54**, 906-910.
- Kennett, B.L.N., and Williamson, P.R., 1987, Subspace methods for large scale nonlinear inversion: Mathematical Geophysics, a survey of recent developments in seismology and geodynamics, edited by N.J. Vlaar, G. Nolet, M.J.R. Wortel, and S.A.P.L. Cloetingh, D. Reidel Co., Dordrecht.
- Koren, Z., Mosegaard, K., Thore P. and Tarantola, A., 1990, Seismic background velocity estimation and model error analysis by simulated annealing, 1990, 52nd E.A.E.G. Meeting.
- Landa, E., Beydoun, W. and Tarantola, A., 1989, Reference velocity model estimation from prestacked waveforms: Coherency optimization by Simulated Annealing, *Geophysics*, **54**, 984-990.

Mosegaard, K. and Tarantola, A., 1990, Inverse Problems and Simulated Annealing, Submitted to *J. Geoph. Res.*

Tarantola A., 1986., A strategy for non linear elastic inversion of seismic reflection data *Geo-*

physics, **51**, 1893-1903.

Tarantola A., 1988. Theoretical background for the inversion of seismic waveforms including elasticity and attenuation, *Pure and applied Geophysics*, **128**, n° 1-2, 365-399.

Figure 1 Shot record 167 after first arrival muting.

Figure 2 P velocity initial model at location 7000 m obtained by Total-CFP using focusing analysis optimization technique.

Figure 3 True velocity modek at location 7000 m .

Figure 4 Top, P impedance map after the first iteration.
Bottom, P impedance map (same as above) with main geological structures and velocity contrasts.

Figure 5 Modeled data after one iteration, which corresponds to same shot point as figure 1.