LA-UR-97-3381

Approved for public release; distribution is unlimited

Title: **Computational 3-D Inversion for Seismic Exploration**

RECEIVED
OCT 01 1997
OSTI

Author(s): E. M. Gavrilov, CIC-ACL
D. W. Forslund, CIC-ACL
M. C. Fehler, EES-4

Submitted to: DOE Office of Scientific and Technical Information (OSTI)

Los Alamos
NATIONAL LABORATORY

DISCLAIMER

Computational 3-D Inversion for Seismic Exploration

Eugene M. Gavrilov* and David W. Forslund
Computing, Information, and Communications Division

Michael C. Fehler
Earth and Environmental Science Division

Abstract

This is the final report of a four-month, Laboratory Directed Research and Development (LDRD) project carried out at the Los Alamos National Laboratory (LANL). There is a great need for a new and effective technology with a wide scope of industrial applications to investigate media internal properties of which can be explored only from the backscattered data. The project was dedicated to the development of a three-dimensional computational inversion tool for seismic exploration. The new computational concept of the inversion algorithm was suggested. The goal of the project was to prove the concept and the practical validity of the algorithm for petroleum exploration.

Background and Research Objectives

Today only a few conceptually different tools are used routinely in seismic exploration. They fall into these groups: deconvolution, stacking, migration and forward modeling. None of the mentioned methods provides adequate inversion without interactive and multi-step corrections and adjustments made by humans ("interpreters"). The ideas used today are fifty years old, they were born when computing was performed with a slide rule. The new generations of supercomputers have brought about new possibilities to explore the large-scale regression methods that require hundreds of macroscopic forward-modeling runs—something that was not possible even five years ago. Although the three-dimensional (3-D) inversion remains a formidable and mathematically ill-posed problem, even for the simplified approximations, the "solution" seems conceivable. The research objective of the project was to prove the concept of the new mathematical approach, test it, and to explore feasibility of its industrial applications. The proposed method could potentially improve the existing technology in the gas and oil industry.

* Principal Investigator, E-mail: gavrilov@lanl.gov

Importance to LANL's Science and Technology Base and National R&D Needs

The crucial stage of the current gas and oil exploration technology is the creation and analysis of seismic surveys. Today, the average price tag for one drill hole starts at $10M. The risk of getting a "dry" hole can be rather high. That is why the companies today spend on the average up to $5M to get the preliminary survey and analysis of the potential field. The huge redundancy of gathered data and the subsequent computer and human analysis of the results explain the high cost of this stage. Still, the accuracy of the current methods cannot be mathematically expressed because it is a human who "draws" the final result, not a computer. Interpreters, not the mathematical algorithm, make the conclusion about the structure of the ground. Clearly, it would have been a tremendous breakthrough if there was a self-consistent tool that would eliminate the human involvement and human errors. It would have been even better if the new methods could also reduce the redundancy of the gathered data thereby increasing production efficiency.

The formidability of the problem comes from the fact that to test an idea one should use 3-D codes and a pretty high resolution. The problem cannot be scaled down to the smaller space dimensions because the fundamental solution of the wave equation is completely different for 1-D, 2-D, and 3-D spaces. At this time, it seems that only the National Laboratories with their great computational power are capable to attack this kind of problem. Just a few major oil companies have comparable computer power, but it is mostly dedicated to the current production regime. The development of a new and better technology for the energy industry is within the Laboratory and DOE missions, and national science interests.

Scientific Approach and Accomplishments

The general statement of the inverse problem can be formulated in a simplified form as follows. There is an investigated nonuniform media. There are sources S (that will send the signals of a given shape into the media) and receivers R on top of the media. It is presumed that everything can be measured only on top of the media (see Fig.1). The problem is to find the properties of the media (in our case the velocity $C(x,y,z)$) from the signals recorded at the receivers (backscattered data).

It is still an open question whether the problem has a unique solution[1] (unique $C(x,y,z)$ will result in a unique pattern recorded by receivers), and the problem is known as ill-posed,[2] i.e., small changes in $C(x,y,z)$ will generate a totally different pattern of receiver records.

Three approaches were considered to attack the inverse problem in the current project. All three methods use scalar 3-D nonuniform wave equation as a forward modeling tool. The code for the 3-D wave equation was developed at LANL, together with the French Petroleum Institute (IFP), under the gas and oil national information infrastructure project. The code runs on a massively parallel computer, Cray T3D, and is capable of simulating realistic problem sizes (up to 10 km^3). The parallel code for the T3D has been implemented at the Advanced Computing Laboratory (ACL) at LANL by one of the authors (Eugene Gavrilov). The code has two types of boundary conditions for the top ground layer, i.e., the absolute reflection to simulate acoustic waves and the transparent boundary conditions to simulate electromagnetic wave propagation. The walls and bottom of the computational box are transparent to simulate the infinite domain. The typical run takes about an hour on the 64-node partition of T3D for one forward simulation. This code was used to check the inverse problem algorithms as well as to construct the regression-like algorithm for the inverse problem. Below we describe briefly the ideas and algorithms used to approach the problem and our accomplishments.

The first method is based on the possibility of extracting information about the nonuniform media from the data recorded on top of the surface with the recurrent chain of functions that relate directly the signal at the receiver with the velocity $C(x,y,z)$. This method is based on the forward model and its discrete approximation. Let us consider the scalar wave equation in acoustic approximation:

$$u_{tt} = c^2\Delta + f(t)\delta(x - x_s) \qquad (1)$$

Here u is the acoustic pressure in the media, $C = C(x,y,z)$ is the velocity in the media, $f(t)$ is the time profile of the source, located at X_S. One of many finite-difference approximations for the equation above can be as follows:

$$u^{n+1}_{i,j,k} - 2u^{n}_{i,j,k} + u^{n-1}_{i,j,k} = c_{i,j,k}{}^2(\tau / h)^2(u^{n}_{i+1,j,k} + u^{n}_{i-1,j,k} + u^{n}_{i,j+1,k} + u^{n}_{i,j-1,k} + u^{n}_{i,j,k+1} + u^{n}_{i,j,k-1} - 6u^{n}_{i,j,k}) + f^n\tau^2 \qquad (2)$$

This equation approximates the original wave equation and will provide the solution with *a priori* chosen accuracy as long as computational parameters, such as the time-step τ and space step h, are chosen correctly to provide stability and convergence of the scheme. But the same equation can be considered as a recurrent functional that maps the value of the pressure $\boldsymbol{u(x,t)}$ at an arbitrary point in space $\boldsymbol{x}$, e.g., on the surface, to the space of $\boldsymbol{C(x,y,z)}$. The initial value problem

$$u(x,t=0)=0 \qquad \textbf{(3)}$$

$$u_t(x,t=0)=0 \qquad \textbf{(4)}$$

makes this functional mapping finite in $C(x,y,z)$ space for the finite time of solution. That means that for an arbitrary source $f(t)$ it is possible to construct the function Φ in such a way that

$$u(x_r,t)=\Phi(c(x)|f(t)) \qquad \textbf{(5)}$$

Hence, the records at the receivers x_r can provide information about the distribution of velocity in the finite domain surrounding the source and receivers (we assume there are no singularities in the media). The problem then becomes how to find $\Phi^{(-1)}$ (inverse to Φ) that will "invert" the records into numerical values of $C(x,y,z)$. One approach that we have chosen was to iteratively solve Eq. (2) with respect to $\boldsymbol{C(x,y,z)}$ using the records $\boldsymbol{u(x_r,t)}$ obtained from forward modeling. The idea was very attractive; however, even for the second-order finite-difference approximation [Eq. (2)], the computational intensity simply did not let us move any further then just a few points from the receiver location, even with the massively parallel computer. The structure of $\Phi^{(-1)}$ is so severe, from the computational point of view, that currently it does not make sense even to speak about the practical value of the approach. All attempts to use the layering approach (linearization) in decomposing Eq. (5) led us to the absolutely unstable algorithms that usually blew up just after a few steps. There was no way to regularize the function $\Phi^{(-1)}$ because it is a nonlinear recurrent mapping and there are no techniques to do this. Although mathematically this approach seemed to be a conceivable way to solve the problem, only negative results were obtained and the authors have to admit that the method simply did not work out.

The second method we tried was based on the idea of regression approximation of the solution $\boldsymbol{u(x,t)}$ via the set of basis functions constructed from the velocities $\boldsymbol{C(x,y,z)}$ and time-dependent coefficients $\beta(t)$. The goal was to find a correlation between particular space distribution of the velocity in the media and the responses that are collected on the ground surface. Again, Eq. (2) was chosen as an approximation of the functional dependencies. It should be pointed out that for this approach, Eq. (1) should be dimensionless and all values normalized with respect to the "natural" or characteristic values. Our normalization was the following:

$$\hat{t} = t\nu \qquad (6)$$

$$\hat{x} = x / \lambda \qquad (7)$$

$$\lambda = c_s / \nu \qquad (8)$$

$$\hat{u} = u / A, \qquad (9)$$

where ν is the central frequency of the source, λ is the characteristic wavelength, c_s is the velocity at the source position, and A is the amplitude of the source. Choosing these natural variables the velocities of the dimensionless wave-equation are expressed in terms of the velocities at the source location, the time is expressed in inverse frequency and the pressure is normalized on the amplitude of the source. The actual approximate relations between u and C were chosen as follows:

$$\hat{u}_r(t^k) \approx (\beta_0(t^k) + \beta_1(t^k)\hat{c}^2{}_1 + \beta_2(t^k)\hat{c}^2{}_2 + \cdots + \beta_m(t^k)\hat{c}^2{}_m) \qquad (10)$$

Here $u_r(t)$ is the recorded pressure at the receiver, located at the same position as the source, $\beta_l(t)$ are the time-dependent coefficients that relate dimensionless $u_r(t)$ and dimensionless velocities c_i in the finite domain. Eq. (10) states that there is an approximate relation between the records and the velocities in linear approximation. If Eq. (10) holds, then the inverse matrix B would immediately provide information about the velocities from the records, i.e.:

$$B = \begin{matrix} \beta_0(t^0) & \beta_1(t^0) & \beta_2(t^0) & \ldots & \beta_m(t^0) \\ \beta_0(t^1) & \beta_1(t^1) & \beta_2(t^1) & \ldots & \beta_m(t^1) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \beta_0(t^m) & \beta_1(t^m) & \beta_2(t^m) & \ldots & \beta_m(t^m) \end{matrix} \qquad (11)$$

$$\bar{c} = \|B^{-1}\| \bar{u} \qquad (12)$$

To obtain the matrix *B* we had to use the Galerkin projection method[3] and made several dozens of forward model runs. In fact, the more runs we have the better we can "memorize" the different velocity structures and the trace record patterns corresponding to these velocities. This is a regression approach. It is also implied that the chosen projection, Eq. (10), will have good generalization ability, i.e., it will produce, with reasonable accuracy, the proper trace records not only for the velocities that we used to "train" it but also for the arbitrary velocity structures. The matrix *B* coefficients are calculated as follows:

$$\langle \hat{u}_r(t^l) | \hat{c}^2{}_k \rangle \approx (\langle \beta_0(t^l) | \hat{c}^2{}_k \rangle + \langle \beta_1(t^l)\hat{c}^2{}_1 | \hat{c}^2{}_k \rangle + \ldots + \langle \beta_m(t^l)\hat{c}^2{}_m | \hat{c}^2{}_k \rangle) \qquad (13)$$

Here the angle brackets mean scalar product (integration over the set of runs). The approach is formidable because to compute matrix *B* one should first have "enough" forward runs to solve the system [Eq. (13)]. Also, the system [Eq. (13)] is a function of time, i.e., it provides a solution only for one row of the matrix *B*. We have to remember that in order to get reasonable accuracy one should have enough time and space resolution, i.e., for parameters of practical interest, e.g., when $\nu \approx 10$ Hz, $c_s \approx 2000$ m/s, $\lambda \approx 200$ m, and domain of interest up to 10 km in each direction, the size of the matrix *B* is large (at least *1000×1000*). It is clear that to compute a "good" *B* one needs at least several dozens of forward runs from which approximately 1000 solutions have to be obtained for the linear systems of the size *1000×1000*. Clearly this problem is for supercomputers. It would have taken about a week of computer time only to accumulate enough runs. However, once computed, matrix *B* can be used to invert the velocities for practical applications over and over again. It would have been even possible to "hardwire" the matrix and the linear Eq. (12) on a PC board to have a real-time inverter.

At this stage we have tried to scale down the problem a bit to see whether the method will provide adequate results, at least approximately. The goal was to test the concept. Due to the very short time allocated for this project, thorough investigation was not possible but some preliminary results are available. We have chosen the following method of finding the matrix *B*. The forward runs were done at the ACL on the Cray T3D. These runs provided the record sets for the corresponding sets of arbitrarily chosen velocities. The decimated velocity grids $\{c_1 ....c_m\}$with corresponding trace records $u(t^l)$ were stored to use the Galerkin method to solve system (13). The systems were solved on the massively parallel computer, CM-5, at the ACL at Los Alamos. This computational "setup" allowed us to speed up the overall computational time and may be considered as our main result at this stage. We also found that, on the small models, Eq. (12) has a reasonable generalization ability, at least for the simple-layered or clustered velocity structures. However, there is a need to investigate this approach further.

The third method used explores the fact that one can use electromagnetic waves for petroleum exploration (from short waves to decimeter diapason waves) (e.g., Ref. 4 and Ref. 5). The fact is that the ground mostly has constant dielectric permeability for a wide range of electromagnetic frequencies, and there is a rather big difference between the ground permeability and permeability of gas and oil type clusters. This fact is being used today by many gas and oil and service companies in their production exploration. There are good methods developed for this approach (frequency domain); see a good survey of these methods in Ref. 6. The third computational method that we have developed neither solves the inverse problem, nor uses a new mathematical idea. However, the computational implementation provides the capability to reconstruct underground images in search for gas and oil and any other conductive deposits (metals). Since for electromagnetic waves the discontinuities of permeability most probably are due to the presence of conductive type structures, one can reconstruct a 3-D image of these clusters. The images of the cluster surfaces will not be optically distorted and can be used to get an estimation of size and location depth of these clusters.

The idea of image reconstruction is the same that is used in holography, i.e., a stationary electromagnetic source sends a continuous signal and a few receivers collect backscattered radiation. The receivers collect amplitude and phase information. This can be done either by applying a reference signal of known frequency or by using a computer post-processing of time-varying receiver signals. Either way we can reconstruct the amplitude and the phase for the surface layer. This information is used as a boundary condition to solve the Helmholtz equation for the 3-D space above the surface layer. The solution is a field generated by scattered centers. The amplitude of the 3-D field is a 3-D image of the clusters.

The Helmholtz equation for the outer space can be solved using Fresnel-Kirchgoff theory or Huygens' principle,[7] i.e., field *u(x)* can be computed as the convolution of the surface field (or field just from few receivers) and Kirchgoff kernel:

$$u(\bar{x}|k) \approx \iint A(\bar{x}')K(\bar{x}-\bar{x}'|k)d\bar{x}', \qquad (14)$$

where

$$K(\bar{x}) = -\frac{ik}{4\pi}(1+\cos\alpha) \qquad (15)$$

Here *k* is the wave number corresponding to the wave number for outer space (for practical applications *k* is just a computational parameter), α is an angle between the surface normal and direction toward $\vec{x}'$, A(x) is the amplitude and phase information (hologram), collected on the surface, K(x) is the Kirchgoff kernel. We implemented integration of Eq. (14) on the Cray-T3D at the ACL. Below we present a few pictures that we have obtained by applying this method to the synthetic data set. The forward model was used to calculate *A(x,y)* for the chosen ***C(x,y,z)***. Then *A(x,y)* was used to reconstruct the image of ***C(x,y,z)***. For this particular example we considered a 3-D case when a source generated electromagnetic waves into the media that had a nonuniform cluster (spherical shape of a given radius) located in the middle of the computational box. Velocity contrast (ratio of velocities for surrounding media and the cluster) was equal to 0.5. There was one source located on the surface and shifted toward the upper left corner of the computational box.

Figure 2 shows the hologram *A(x,y)* (amplitude-phase information) recorded on the surface from forward modeling. This hologram was used to reconstruct the wave field, i.e., to invert the wave-front using Eq. (14). From the inverted wave-field *u(x,y,z)* we have calculated the amplitude $u^2(x,y,z)$. Figures 3(a) and 3(b) show the vertical slice across the center of the 3-D computational box: Fig. 3(a) - the actual velocity structure, Fig. 3(b) - the reconstructed amplitude that shows our cluster. Fig. 4(a) shows the vertical slice of the 3-D velocity model, slightly shifted, compared to Fig. 3(a). Fig. 4(b) shows the vertical slice, same location as in Fig. 4(a), of the 3-D reconstructed amplitude.

The results show that the method has great potential for practical applications but, again, further analysis and optimization are needed to convert it from a working idea to a practical tool.

References

1. Rakesh And W. Symes, "Uniqueness for an Inverse Problem for Wave Equation," *Commun. in PDE* **13**, 87–96, (1988).

2. M.Lavrentiev, V. G. Romanov, and S. P. Shishatski, "Ill-Posed Problems of Mathematical Physics and Analysis," American Mathematical Society, Providence, Rhode Island, (1986).

3. Albert Tarantola, "Inverse Problem Theory," Elsevier Science B.V., Amsterdam, the Netherlands, (1987).

4. M. N. Berdichevsky, and M. S. Zhdanov, "Advanced Theory of Deep Geomagnetic Sounding," Elsevier, (1984).

5. C. Degroot-Hedlin and S.C.Constable, "Occam's Inversion to Generate Smooth Two-Dimensional Models from Magnetotelluric Data," Geophysics, **55**, 613–1624, (1990).

6. D. Colton and R. Kress, "Inverse Acoustic and Electromagnetic Scattering Theory," Springer-Verlag, New York, (1992).

7. T. S. Yu Francis, "Introduction to Diffraction, Information Processing, and Holography," MIT Press, (1973).

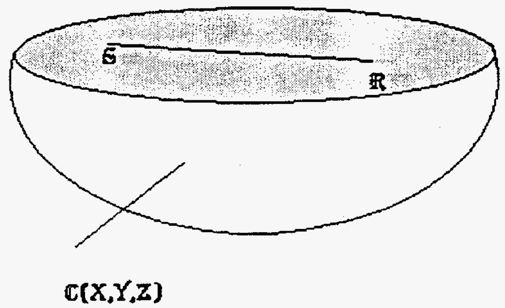


Fig. 1. Inverse problem spatial layout. ***C(x,y,z)***–Velocity to be determined; ***S***–Source; ***R***–Receivers.

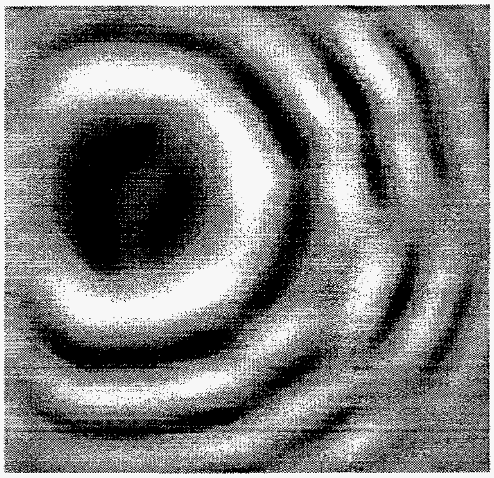
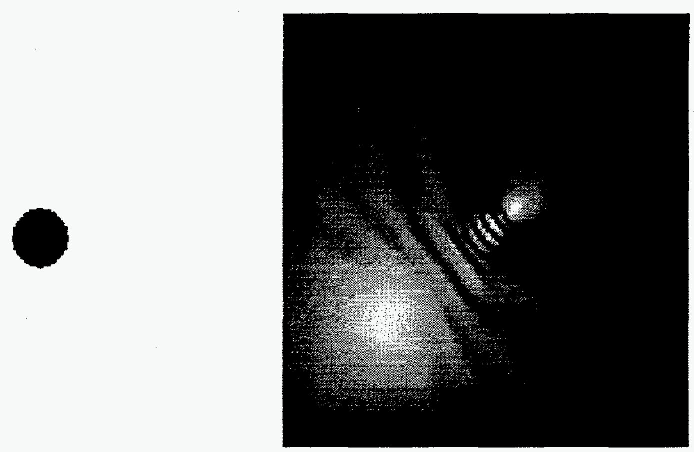


Fig. 2. Hologram *A(x,y)* from forward modeling.

(a) (b)

Fig. 3. (a)Vertical slice of velocity across the center, Synthetic Model. (b)Vertical slice of reconstructed amplitude across the center of the model.



(a) (b)

Fig. 4. (a)Vertical slice of velocity slightly shifted from the center position. (b)Vertical slice of reconstructed amplitude for the same spatial location as in Fig. 4(a).