Differential Semblance Estimation of Velocities: Progress Report

William. W. Symes *

ABSTRACT

Recent progress and plans for future work on differential semblance optimization center around multiple dimensions, multiple reflections, and multiple refractions.

Introduction

Velocity estimation is based on semblance. Measuring the semblance differentially - comparing images with nearby acquisition parameters - offers several mathematical and practical advantage. For an account of these, the reader can consult (Claerbout, 1992), pp. 93 ff, as well as previous papers of the author (Symes, 1994).

In several recent reports I have documented a version of differential semblance velocty estimation based on layered modeling of the subsurface and high frequency asymptotics (Araya et al., 1996; Symes, 1997). This version ("CMP/NMO") is computationally efficient, flattening a field CMP with 100 3 s traces in a few seconds on a modest workstation. However its scope is limited: because of the modeling assumptions on which it is based, it accomodates only very mild lateral velocity variations and low levels of coherent noise such as multiple reflections. Complex refraction effects such as occur around and below salt and fault block structures are as far beyond its ken as they are beyond ordinary NMO-based velocity analysis, for the same reasons. Because of implementation details it requires essentially regular 2D source - receiver geometry. This report describes an "upgrade path" intended to remove or diminish these limitations, and recent progress in accomplishing the intended modifications.

*The Rice Inversion Project, Department of Computational and Applied Mathematics, Rice University, Houston TX 77251-1892 USA, email `symes@caam.rice.edu`

Multiple Dimensions

Estimation of velocities with mild lateral variation from multioffset primary reflections is in principle no more difficult than estimation of layered models, but demands computational efficiency. There is nothing new about estimating laterally heterogeneous velocity models by DSO - the 1993 and 1994 TRIP Annual Reports contain several examples in which DSO successfully recovers models with significant lateral heterogeneity from synthetic multioffset data. However these excercises employed finite difference full waveform modeling as the driving simulator. The resulting code was simply too CPU-intensive to serve as a tool for analysis of industrial scale data, or even as a platform for serious exploration of the DSO algorithm. Moreover in the meantime the theory of DSO evolved through the theses of Hua Song and Mark Gockenbach (Song, 1994; Gockenbach, 1994), and the theoretical developments suggested interesting algorithmic refinements.

One obvious option for enhanced efficiency was to replace the finite difference full waveform simulation by high frequency asymptotics, and that is the route we have pursued at TRIP. We are no longer alone in this: the thesis work of Hervé Chauris, in the Ecolé des Mines group led by Pascal Podvin and Gilles Lambaré, explores multidimensional DSO using the HFA tools developed by that group (Chauris et al., 1998a; Chauris et al., 1998b), with very positive results. We have used the Kirchhoff modeling, migration, and inversion package developed at TRIP (Symes et al., 1994; Araya and Symes, 1996), based on high order paraxial (limited aperture, space marching) finite difference schemes for traveltime and amplitude, also developed here. Work on numerical kinematics continues at TRIP, as evidenced by other papers in this report, and will eventually find its way into DSO implementations.

So far as I know, another obvious option has not been explored: no reports have appeared to my knowledge of DSO based on paraxial wave simulation.

The general framework for "raw" DSO is simple. It relies on a binning scheme for seismic images: common shot and common offset are the two obvious examples. If d denotes multioffset data, v a velocity model, $F[v]$ a (linearized) prestack forward modeling operator, $G[v]$ a prestack linearized inversion operator approximately inverting $F[v]$, and D a difference operator expressing the differences between *neighboring* image bins, then the raw DSO objective function is

$$J_0[v;d] = \frac{1}{2}\sum |F[v]DG[v]d|^2$$

where the sum is over all traces and time samples. The step $d \rightarrow G[v]d$ constructs an amplitude-corrected image, the application of D measures semblance "differentially", and the application of $F[v]$ suppresses the effect of amplitude variations. Theoretically, D should include a factor which controls the growth of high frequency components, but numerical experiments suggest tentatively that one can ignore this requirement with some impunity, and use for D a simple image gather - to - image gather difference operator.

For a description of "cooked" DSO, see the section on multiple reflections below.

In the current "2.5D" implementation, $F[v]$ and $G[v]$ are the high frequency asymptotic modeling and inversion operators described in (Symes et al., 1994; Araya and Symes, 1996). At the moment the code implements the function value. The gradient is necessary for optimization, and is in the design stage. Some component computations of the gradient are discussed below.

Note that because the TRIP DSO code follows an object oriented design, the 2.5D package differs from the 1.5D package only in the code implementing the operators specific to multimensional scattering. The DSO calculations themselves are common to all instances. Because of this design feature, the transition from 1.5D to 2.5D (sans gradient) took less than a day.

Figure 1 shows a common midpoint gather from a marine data subset donated by Shell Research. Preprocessing consisted of low-pass filter and mute; I also applied predictive deconvolution to suppress an obvious strong surface-related multiple train, traces of which are still evident. A near offset section from this line appears as Figure 2. Application of CMP/NMO DSO to the CMP produced an estimated interval velocity profile; the related RMS velocity is displayed in Figure 3. The quality of this estimate will be discussed in the next section. Note that the common offset section suggests that mostly flat structure exists down to perhaps two seconds in this part of the North Sea, so that CMP/NMO DSO should produce relatively accurate velocities, so long as other underlying conditions of its validity are met.

Figure 4 shows a number of values of J_0 computed along the line segment in layered velocity model space between nominal water velocity and a velocity profile similar to but larger than that in Figure 3, using the 2.5 D Kirchhoff based DSO. For this calculation I used a subset of the Shell data consisting of every other offet and every other shot. This amounted to 120 offset gathers each containing 100 traces. I used offset as the bin parameter, so $G[v]$ for example is common offset inversion (Araya and Symes, 1996). Each function value required approximately 8 hours of CPU on an Intel Pentium II 300 MHz machine running FreeBSD 2.6, Gnu C and C++ compilers, and `f2c` to compile the Kirchhoff code which is entirely in Fortran. I believe that this performance is suboptimal, and that considerable improvement is attainable.

The minimum along this line segment occurs at approximately 80% of the target velocity, which was itself nearly 120% of the DSO velocity estimate.

Low level oscillation is visible in the values of this linear slice of J_0, whereas the theory assures us that J_0 is smooth. I believe that spatial aliasing is responsible for this effect. An effective antialiasing criterion for DSO computations is clearly a prime necessity, and is high on my list of future projects.

Actual velocity inversion will require a gradient. To see what is required, consider the integral form of $F[v]$ (Beylkin, 1985; Symes et al., 1994):

$$F[v]r(t) = \int dx\, A(x)\delta(t - \tau[v](x))$$

in which r is the vector of relative Earth model parameter fluctuations, A is the amplitude which depends linearly on r, and $\tau(x)$ is the two way time from source to subsurface point x to receiver in the velocity field v. All three of these functios also depend on source and receiver coordinates.

The rate of change of $F[v]r$ with respect to v in response to velocity perturbation δv is

$$D(F[v]r)[\delta v](t) = -\frac{\partial}{\partial t}\int dx\, A(x)\delta(t - \tau[v](x))D\tau[v][\delta v](x)$$

up to lower order terms of the sort neglected in this theory. From this expression one sees that the adjoint of the linear map $\delta v \mapsto D(F[v]r)[\delta v]$, which is the critical component in the gradient of $J_0[v]$, involves the adjoint of the traveltime derivative $\delta v \mapsto D\tau[v][\delta v]$. This latter adjoint is exactly the backprojection operator of traveltime tomography.

In (Symes and Sei, 1994; Symes and Sei, 1995) Alain Sei and I showed how to calculate the backprojection operator using the adjoint state technique and paraxial finite difference eikonal solvers, i.e. without tracing any rays, computing a path length matrix, or any of the other usual paraphenalia of ray based tomography. It is necessary to combine this calculation with the high frequency asymptotic integral representation of F (and the similar representation of G) to produce the operator velocity derivatives and their adjoints. Once this is done, there is essentially no additional work required to compute the gradient of J_0, due once again to the object oriented design of TRIP DSO.

Acquisition Geometry

Even when the earth is close to layered, so that the 1.5 D model should be appropriate, field multioffset data may differ in a very important respect from the ideal underlying the current CMP/NMO DSO implementation: the sampling in source and receiver coordinates may not be regular.

To show how the assumption of regular sampling may affect DSO, I display in Figure 5 a CMP from a marine line contributed by Mobil Upstream Strategic Research Center, and in 6 a NMO corrected ("prestack inverted") CMP using the optimal velocity computed by DSO. The only preprocessing consisted of low pass filter and mute. The event at 1.6 s exhibits considerable residual curvature, and no amount of tuning or parameter fiddling produced an appreciably better result.

This "failure to flatten" is due entirely to the assumption of regular sampling. In preparing the data for DSO, I had computed an offset interval by averaging offset increments from the trace headers, and used this assumed regular sampling in all subsequent processing. However a plot of the source and receiver positions from the trace headers, Figure 7, suggests that the offset interval changes systematically with trace number, because the lines along which source and receiver points lie are not parallel! This happens for a very good reason: the data was shot with a two boats in order to undershoot a platform, and this particular midpoint is near the platform around which the boats had to maneuver. Assuming regular sampling in this situation is enough to prevent proper moveout analysis by DSO or any other means.

The solution is clearly to use the trace header geometry, which while not perfectly reliable is still the best information on acquisition geometry available to the velocity estimation process. The S Vector class, described elsewhere in this report, retains all trace header information and fits into the object oriented design of DSO and other TRIP software. All of the TRIP packages, including DSO, will use S Vector as their basic data vector class in their next editions. At that point our software will take account of actual sampling, insofar as it is represented in the input trace headers, and the difficulty outlined in this section should largely disappear.

Multiple Reflections

Since the current theory has the same basis as seismic migration, it treats multiple reflections as noise. The DSO velocity estimate is not much influenced by random noise; coherent noise however creates velocity errors as documented in last year's report (see also (Symes, 1994) for a related example). This is inevitable: the algorithm attempts to account for all moveout present in the data. If the noise also exhibits moveout, then the algorithm will compromise between velocities explaining kinematics of signal and noise. In the long term a theory accounting for multiple reflections as signal is very appealing. This section explains an approach to making primary reflection based DSO less sensitive to coherent noise, by an extension of the basic algorithm.

A nice illustration of the effect of coherent noise on DSO velocity estimation comes from the data donated by Shell, mentioned above. Figure 8 is the same CMP as displayed above, but without predictive deconvolution. The event around 1.5 s appears "double": the later part is actually a surface multiple of the strong event just after 0.75 s. [The more or less successful suppression of this multiple by predictive deconvolution is thus essentially an accident!] Application of CMP/NMO DSO produced an interval velocity estimate, the corresponding RMS velocity being shown in Figure 9. The NMO-corrected data appears as Figure 10. Comparison with Figures 3 and 11 respectively shows that the presence of the multiple reflection in this data has caused the RMS velocity to be dragged partway towards the moveout velocity of the multiple event. This is characteristic of the influence of multiples in more or less flat structures.

Gockenbach's thesis and subsequent work (Gockenbach, 1994; Gockenbach et al., 1995;

Gockenbach and Symes, 1997) suggests a remedy for this distressing trend, which will function as a supplement to (rather than a substitute for) multiple suppression techniques like predictive deconvolution. The problem with DSO in the presence of coherent noise is that the algorithm tries to fit the noise moveout as well as the signal (primary) moveout. An obvious solution is to permit noise components to be removed before fitting:

$$J_\epsilon[v;d] \equiv \min_{\delta d}\frac{1}{2}\sum |F[v]DG[v](d+\delta d)|^2 \text{ subj } \|\delta d\| \leq \epsilon\|d\|$$

(the vertical bars denote RMS power). Notice that when $\epsilon = 1$ you can chose $\delta d = -d$ so $J_1[v;d] = 0$ always. Also note that when $\epsilon = 0$ you recover the "raw" differential semblance function $J_0[v;d]$ as defined above. So the parameter ϵ naturally varies from 0 to 1.

If coherent noise is present in the data with moveout different from that of the signal, then $J_0[v;d] > 0$ for any velocity v, because it's impossible to fully moveout correct the data. Write d_0 for the "signal", i.e. the part of the data regarded as primary with a definite moveout, and n for the noise component; that is $d = d_0 + n$. The for $\epsilon = \|n\|/\|d\|$, there must be a velocity v_0 for which $J_\epsilon[v_0, d] = 0$ - namely the velocity with which the primary data d_0 moves out! In effect, choosing this value of ϵ permits you to just remove the noise from the data in the definition of J_ϵ, ie. choose $\delta d = -n$.

It may be that the noise has a definite moveout corresponding to another velocity. This would be the case for example for a single family of surface related multiples. Then you could also achieve $J_\epsilon[v,d] = 0$ by choosing $\epsilon = 1 - \|n\|/\|d\|$, taking v to be the moveout velocity of the noise, and $\delta d = d_0$ - ie. remove the primary signal and flatten the multiple!

The key observation is this: if the primary power is greater than the power in the coherent noise, then the second alternative requires a larger ϵ than the first.

Define the *semblance error function* $H(\epsilon)$ by

$$H(\epsilon) = \min_v J_\epsilon[v;d]$$

i.e. for each ϵ find the velocity model and data perturbation of size at most ϵ with the smallest differential semblance. Then the observation above suggests the *noise-finding algorithm* for DSO velocity analysis:

Find the smallest zero of $H(\epsilon)$, and you have found both the primary moveout velocity and the noise power!

I emphasize that this concept is appropriate when the (coherent) noise is smaller than the signal, or has been suppressed. This should apply to high quality reflection data that has been processed to remove multiple reflections by one of several modern techniques. The point is that it is not necessary to remove multiples - only to suppress them to the point that they have substantially less power than the primaries, without altering the primary signal.

The results of (Gockenbach, 1994; Gockenbach et al., 1995; Gockenbach and Symes, 1997) show that $H(\epsilon)$ is convex and give a means to compute its derivative - in fact the derivative is just a by-product of the computation of $J_\epsilon[v;d]$ at the optimal v. Thus approximating the first zero fo $H(\epsilon)$ is rather easy (see the picture on p. 60 of (Gockenbach and Symes, 1997)). Also the gradient of $J_\epsilon[v,d]$ is computed by almost exactly the same steps as the gradient of $J_0[v;d]$.

So the key step in finding the first zero of $H(\epsilon)$ is the solution of the problem defining J_ϵ. This involves the minimization of a quadratic objective with a quadratic constraint; numerical analysts know this as "the trust region problem". The efficiency with which this step is accomplished will determine the efficiency of the noise finding algorithm. Mark Gockenback and I are looking into a new and very promising algorithm which reduces the trust region problem to an eigenvalue problem, and solves that by an efficient Lanczos iteration (Santos and Sorensen, 1995; Rojas, 1998).

Multiple Refraction

Multiple refraction or multipathing undermines the semblance principle of primary reflections, on which all velocity analysis is based, as Cliff Nolan showed in his thesis (Nolan, 1997; Nolan and Symes, 1997; Nolan and Symes, 1996). The only way out of this conundrum is to replace the global problem with its complex wave front by a collection of local problems with simple wavefronts. This idea is closely related to Dimitri Bevc's semirecursive Kirchhoff migration (Bevc, 1997) and to the slowness matching algorithm for multiarrival traveltime computations (Symes, 1996).

The easiest version of this approach to explain is the paraxial, in which all incident rays are assumed to be downgoing. Then one divides the subsurface into layers, and redatums to the top of each layer using the current velocity model. Equivalently one can construct areal source field which produce simple incident fronts at the top of each layer. Then one applies the differential semblance criterion to the image produced in each layer. When the redatumed data is actually (close to) that produced by simple incident wavefronts, the semblance principle is revived and drives velocity estimates in and above the layer. This will occur as soon as the velocity is not too greatly in error in and above the layer. In the near-surface, this latter condition holds *ab initio*. As shallower layers are refined, deeper layers begin to generate useful semblance information, which also helps to further constrain the shallow model.

The critical step in this algorithm is redatuming. As was the case with Bevc's semirecursive Kirchhoff migration technique and the areal source construction of Berkhout, this is clearly possible in 2D but not at all obvious for typical sparsely sampled 3D data. Cliff Nolan's main thesis result helps clarify the conditions under which differential semblance can be localized as outlined in the last paragraph: these are the same conditions under which imaging can be accomplished. In brief, the projections of incident and reflected rays onto the source - receiver manifold must determine the rays uniquely. See Cliff's publications for further discussion (Nolan, 1997; Nolan and Symes, 1997; Nolan and Symes, 1996).

When refraction is so strong as to illuminate interesting reflectors from below, this "downward continuation" approach will need analogous modifications to those required by the slowness matching algorithm in this case: the subsurface is divided into blocks, and the blocks are visited in roughly the order followed by the rays.

Note that the algorithm sketched here differs from standard velocity model construction in that shallow layers are not fixed early on, but remain in play throughout the iteration.

Summary

Differential semblance applies to multidimensional velocity estimation without any essential conceptual change. Kirchhoff modeling and inversion provide sufficiently fast drivers to make automated DSO velocity analysis of 2D lines entirely feasible on contemporary workstations, and treatment of 3D data sets at least conceivable. An object oriented infrastructure for irregularly sampled data is under construction, and will serve as the foundation for the next generation of DSO implementations.

Two recent TRIP theses are the source of algorithmic attacks on multiple reflection and multiple refraction, respectively. Mark Gockenbach showed how to remove noise systematically from data during velocity analysis, extracting the most coherent signal - which is the primary reflection seismogram in many cases. The implementation depends critically on the solution of the so called Trust Region problem (minimizing a quadratic subject to quadratic constraints), and we are looking into the use of a new and very promising family of algorithms for this subproblem, due to Sorensen and his coworkers. Cliff Nolan's thesis defined necessary and sufficient conditions for imageability of reflector structure from primary reflection data. These are the same conditions under which redatuming can succeed, and with redatuming comes localization of differential semblance and a possible approach to velocity estimation in complex refracting media.

REFERENCES

Araya, K., Kim, S., Nolan, C. J., and Symes, W. W. (1996). Velocity analysis and high frequency asymptotics. Annual report, The Rice Inversion Project.

Araya, K. and Symes, W. W. (1996). 2D and 2.5D kirchhoff inversion using upwind finite difference amplitudes. In *Proc. 66th Annual International Meeting*, pages 503–506, Tulsa. Society of Exploration Geophysicists. Expanded abstract.

Bevc, D. (1997). Imaging complex structures with semirecursive kirchhoff migration. *Geophysics*, 62:577–588.

Beylkin, G. (1985). Imaging of discontinuities in the inverse scattering problem by inversion of a causal generalized radon transform. *J. Math. Phys.*, 26:99–108.

Chauris, H., Noble, M., and Podvin, P. (1998a). Testing the behaviour of differential semblance for velocity optimization. In *Proceedings*, Leipzig. EAGE.

Chauris, H., Noble, M., and Podvin, P. (1998b). Testing the behaviour of differential semblance for velocity optimization. In *Expanded Abstracts*, New Orleans. SEG International Meeting and Exposition.

Claerbout, J. F. (1992). *Earth Sounding Analysis: Processing versus Inversion.* Blackwell, Boston.

Gockenbach, M. (1994). *An abstract analysis of differential semblance optimization.* PhD thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A.

Gockenbach, M. and Symes, W. (1997). Duality for inverse problems in wave propagation. In Biegler, L., Coleman, T., Santosa, F., and Conn, A., editors, *Large Scale Optimization*, New York. Institute for Mathematics and its Applications, Springer Verlag.

Gockenbach, M., Symes, W., and Tapia, R. (1995). The dual regularization approach to seismic velocity inversion. *Inverse Problems*, 11(3):501–531.

Nolan, C. and Symes, W. (1997). Global solution of a linearized inverse problem for the wave equation. *Commun. in PDE*, 22(5&6):919–952.

Nolan, C. J. (1997). *Global analysis of linearized inversion for the acoustic wave equation.* PhD thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A.

Nolan, C. J. and Symes, W. W. (1996). Imaging and conherency in complex structure. In *Proc. 66th Annual International Meeting*, pages 359–363, Tulsa. Society of Exploration Geophysicists. Expanded abstract.

Rojas, M. (1998). *A large scale trust region approach to the regularization of discrete ill posed problems.* PhD thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A.

Santos, S. A. and Sorensen, D. C. (1995). A new matrix-free algorithm for the large-scale trust-region subproblem. Technical report, Technical report TR95-20, Department of Computational and Applied Mathematics, Rice University, Houston, TX 77251-1892.

Song, H. (1994). *On a transmission inverse problem.* PhD thesis, Computational and Applied Mathematics Department, Rice University, Houston, Texas, U.S.A.

Symes, W. (1994). The plane wave detection problem. *Inverse Problems*, 10:1361–1391.

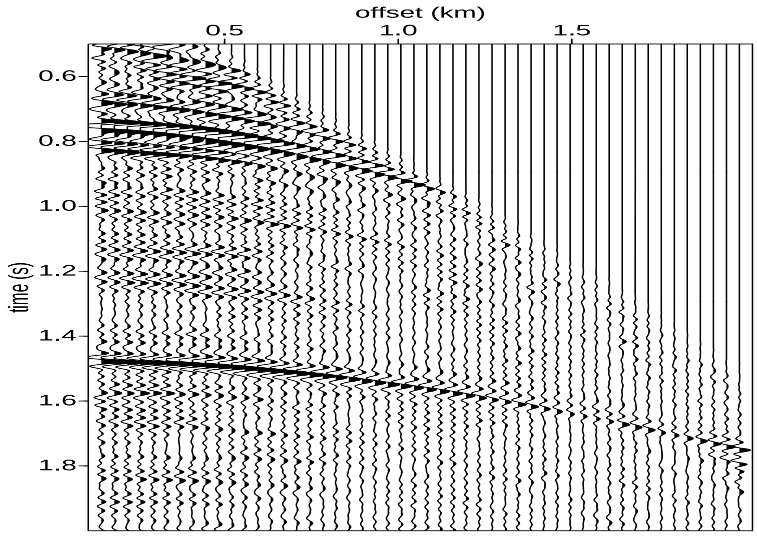

FIG. 1. CMP from North Sea survey line donated by Shell Research. Low pass filter, mute, and predictive deconvolution applied.

Symes, W. and Sei, A. (1994). Gradient calculation of the traveltime cost function without ray tracing. In *Proc. 64th Annual International Meeting of the Society of Exploration Geophysicists*, Los Angeles, CA.

Symes, W. and Sei, A. (1995). Convergent finite-difference traveltime gradient for tomography. In *Proc. 65th Annual International Meeting of the Society of Exploration geophysicists*, Houston, TX. Extended abstract. To appear.

Symes, W., Versteeg, R., Sei, A., and Tran, Q.-H. (1994). Kirchhoff simulation, migration and inversion using finite difference traveltimes and amplitudes. Annual report, The Rice Inversion Project.

Symes, W. W. (1996). A slowness matching finite difference method for traveltimes beyond transmission caustics. In *Mathematical Methods in Geophysical Imaging*. Society of Photooptical Instrumentation Engineers.

Symes, W. W. (1997). High frequency asymptotics and velocity estimation. In *Depth Imaging Symposium*, Tulsa. Geophysical Society of Tulsa.

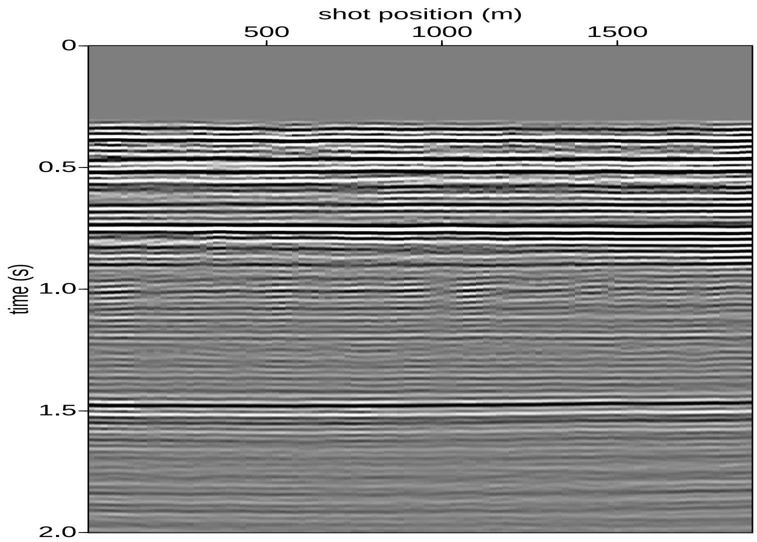

FIG. 2. Near offset gather from North Sea line.

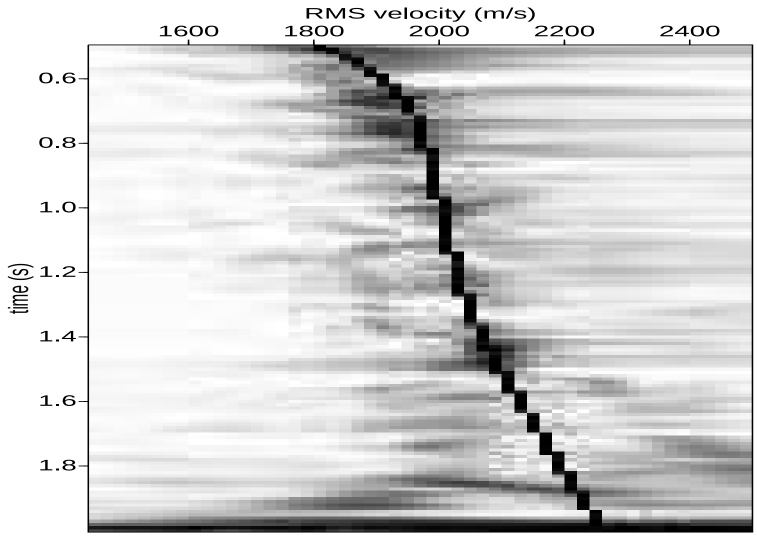

FIG. 3. DSO RMS velocity estimate, overplotted on standard velocity spectrum, for CMP displayed in Figure 1.

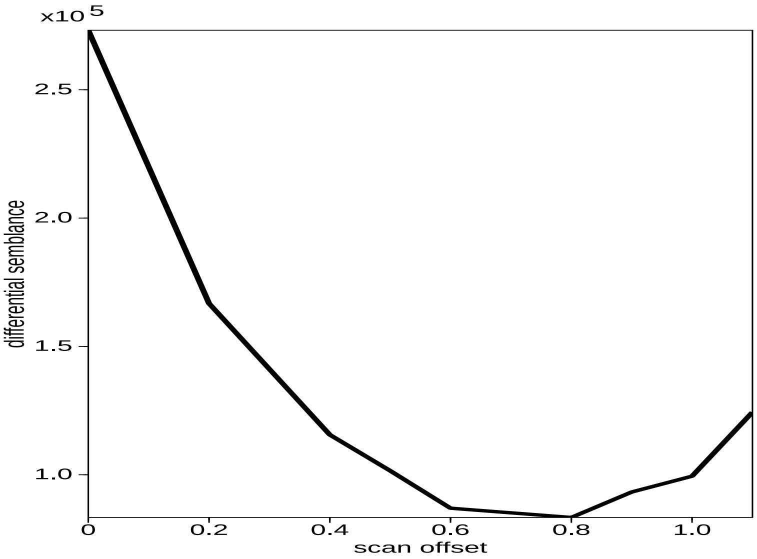

FIG. 4. 2.5D Kirchhoff DSO values from water velocity to another velocity roughly 120% of the CMP/NMO DSO estimate.

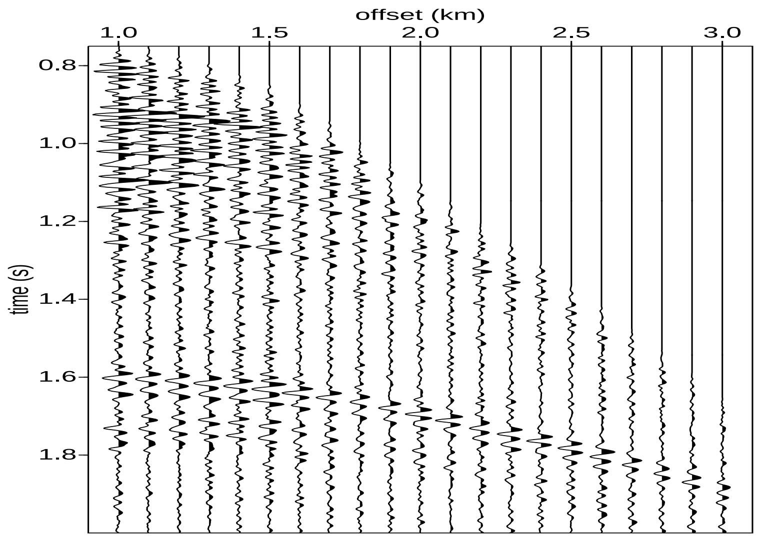

FIG. 5. CMP from marine survey, offshore Nigeria, donated by Mobil Strategic Upstream Research Center. Low pass filter and mute applied.

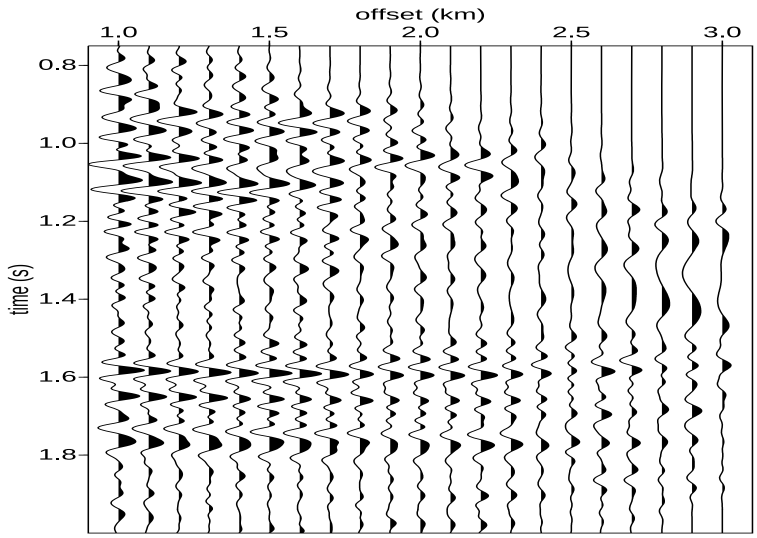

FIG. 6. NMO correction of CMP in Figure 5, optimal CMP/NMO DSO velocity. Note sag in event at 1.6 s.

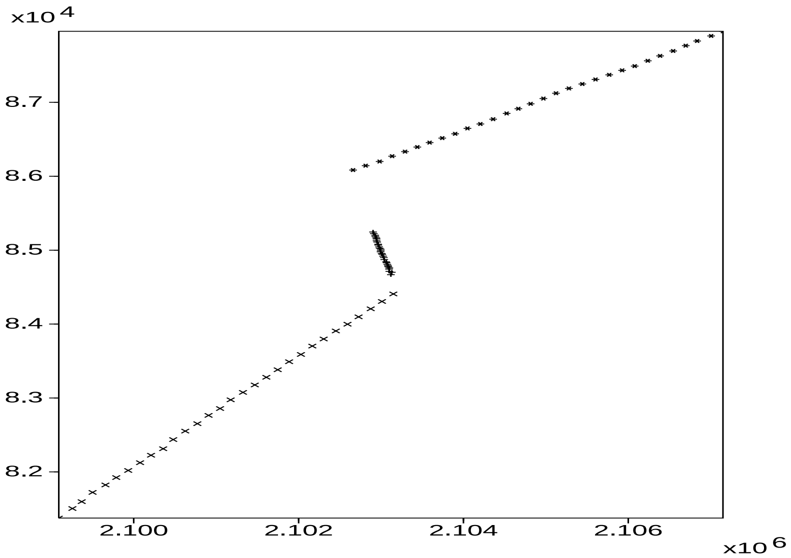

FIG. 7. XY plot of shot positions (x), receiver positions (+), and midpoints (*) for traces in Figure 5. Note that offset spacing is far from uniform - it varies by nearly 40% from near to far offset.

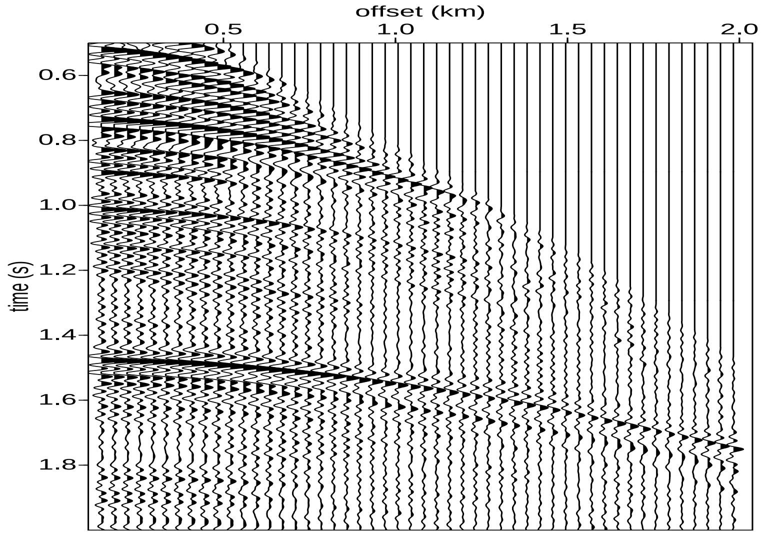

FIG. 8. CMP from North Sea survey line donated by Shell Research. Low pass filter and mute applied.

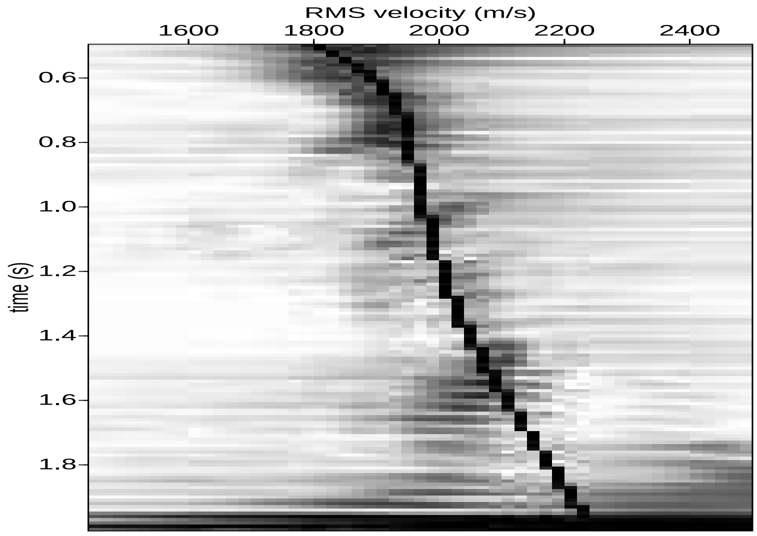

FIG. 9. DSO RMS velocity estimate, overplotted on standard velocity spectrum, for CMP displayed in Figure 8.

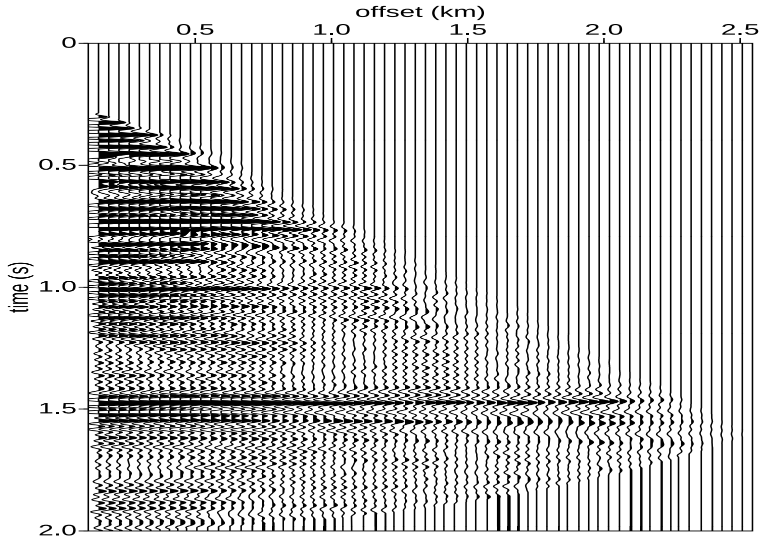

FIG. 10. NMO corrected CMP from Figure 8, using optimal DSO velocity estimate.

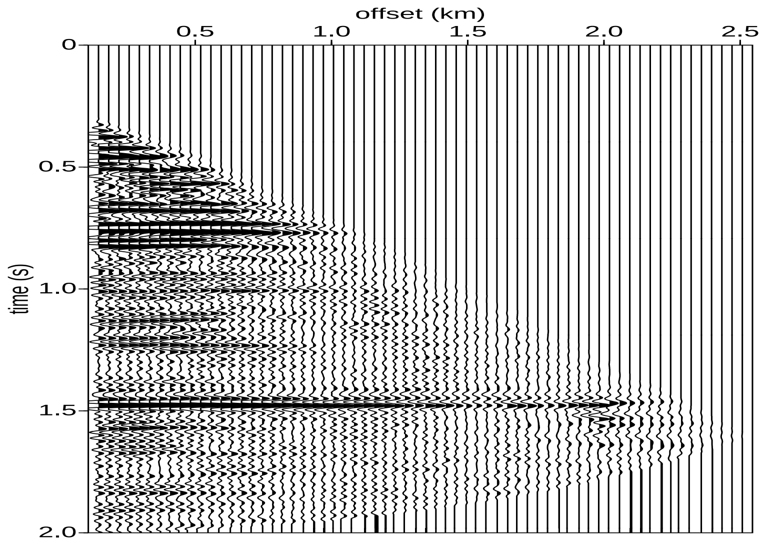

FIG. 11. NMO corrected CMP from Figure 1, using optimal DSO velocity estimate.