Velocity Inversion by Differential Semblance Optimization for 2D Common Source Data

William W. Symes
Michel Kern

April, 1992

TR92-16

Velocity Inversion by Differential Semblance Optimization for 2D Common Source Data *

William W. Symes Michel Kern

The Rice Inversion Project

April 1992

Summary

Differential Semblance Optimization ("DSO") is a variant of data-fitting (least-squares) inversion of reflection seismograms. The misfit functions used in other implementations of least-squares inversion (for example Tarantola 1986, Kolb et al. 1986) exhibit highly nonconvex dependence on *velocity trends*. Therefore least-squares inversion appears to require the use of relatively costly global optimization algorithms such as Monte-Carlo minimization, simulated annealing, or genetic algorithms (Sen and Stoffa 1991, Scales et al. 1991, Tarantola 1991). Also local sensitivity analysis (eg. singular value decomposition of the Hessian) does not describe the resolution limits inherent in such misfit functions - they are too nonlinear. In contrast, the DSO misfit function is smooth and convex over a wide range of velocity models, and so can be minimized satisfactorily *via* efficient local (Newton-like) algorithms. Moreover the quadratic model of the DSO misfit function at the optimum is descriptive of its local behaviour, and so may be used for studies of velocity resolution.

*Expanded abstract submitted to 62nd Annual International Meeting, Society of Exploration Geophysicists.

In previous work (Symes and Carazzone 1991 and references cited there) we have presented DSO inversion for plane-wave data sets and layered media, including application to field data. In this paper, we describe the DSO misfit function for 2D shot gather inversion of laterally heterogeneous models, and an algorithm for minimizing it, and present a simple, preliminary example of shot-gather velocity inversion.

DSO for 2D primaries-only acoustics

The DSO implementation described here is based on the *2D primaries-only constant-density acoustic model*, described by the velocity field v (assumed to be smoothly varying), and the reflectivity field r (assumed to be oscillatory). Both v and r are functions of the spatial coordinates. In shot-gather DSO, the reflectivity is also allowed to depend on the shot coordinate x_s. Since the model must ultimately be independent of shot parameters ("there is only one earth"), this dependence is penalized in the misfit function. The degree to which the inverted reflectivities actually depend on shot parameters is diagnostic of the correctness of the velocity estimate, just as shot-independence (flatness) of common image gathers is diagnostic of velocity correctness in migration velocity analysis (see e.g. Versteeg and Grau 1991).

The predicted seismic data set from v and r is denoted $S[v]r$, and is a sampling at the receiver coordinates of x_r of the primaries-only acoustic field. See e.g. Lailly 1983 or Symes 1991 for a complete description of the boundary value problem defining this field.

The DSO misfit function is a modification of the least-squares error relative to a measured data set S_{data} as follows:

$$J[v] = \min_r \frac{1}{2} \left\{ ||S[v]r - S_{\text{data}}||^2 + \sigma^2 ||\frac{\partial r}{\partial x_s}||^2 + \lambda^2 ||Wr||^2 \right\},$$

where $||\ ||$ denotes the L^2 norm, and W is chosen to be a suitable positive-definite symmetric operator, see Symes 1991.

The roles of the terms in this formula are as follows:

- the first term forces some level of fit-to-data on the choice of r;
- the third term damps r, preventing to some extent the matching of noise in the data;

- the second (differential semblance) term, while also acting to damp incoherent noise in r, has its principal influence on the choice of v: in view of the first term, the second term can only be made small when v correctly predicts the data kinematics.

The second term measures the extent to which neighboring shot-gather inversions for r fail to be the same. Since we measure the misfit between *neighboring* reflectivities, the level of misfit will be indicative of velocity error even when the current velocity estimate is considerably erroneous (i.e. no cycle-skipping occurs). Since the reflectivity estimates are *inversions* rather than merely images, in principle they should be *exactly the same*, rather than merely similar, so that the straightforward difference $\partial r/\partial x_s$ is an adequate measure of semblance.

Gradient calculation

The principal ingredient in local or Newton-type optimization is the calculation of the gradient of J. Note that J is itself defined as the solution of a minimization problem, albeit a quadratic one. It follows from the normal (Euler-Lagrange) equation for this inner optimization (over r) that

$$gradJ[v] = A[r, S[v]r - S_{data}]$$

where

$$< \delta v, A[r, \phi] > = < D_v S[v][\delta v, r], \phi >$$

for arbitrary δv, r, and ϕ; here $<,>$ denotes the L^2 inner product.

In Kern and Symes 1992, we show how to compute the bilinear operator A, by an extension of the now classical adjoint state technique [5]. Once J is known (this involves solving the inner problem for r by a conjugate gradient method), computation of its gradient is possible at a cost of roughly 6 additional forward models. The sequence of operations required is as follows:

1. forward model and linearized forward model;

2. migration (adjoint linearized model) and secondary migration, the latter bearing roughly the same relation to the former as linearized modeling to forward modeling;

3. cross-correlate the forward and linearized forward fields with the adjoint and secondary adjoint fields, to produce a raw before-stack gradient;

4. stack to produce a raw gradient;

5. project into the space of smooth velocity models to obtain the gradient.

In practice, steps 2. and 3. are carried out concurrently; step 3. resembles the imaging step in ordinary before-stack migration. The process is very close to that developed by Chavent and Jacewitz 1990 to compute the gradient of a functional of the migrated image.

The accuracy of the computed gradient depends on the accuracy with which the inner (estimation of r) problem is solved. In fact the formula above is exactly correct only when this inner problem is solved without error. In Kern and Symes 1992 we explain a number of modifications to the calculation outlined above, which increase its accuracy in the presence of inexact estimation of r. All of these enhancments cost CPU cycles, and in the experiments reported here we have used the simple approximation given above.

Nonlinear conjugate gradient iteration

We embedded the gradient calculation of the previous paragraph in the *nonlinear conjugate gradient method* as described in Fletcher 1980, Ch. 4:

1. Initiate:
$$v_1 = v_{initial}, s_0 = 0, \beta_0 = 0$$

2. for k = 1 until convergence do:
$$s_k = -gradJ[v_k] + \beta_{k-1} s_{k-1}$$
$$v_{k+1} = v_k + \alpha_k s_k$$

The line search step α_k is chosen by a backtracking strategy (Fletcher 1980, pp. 26 - 27), and the direction update parameter β_k by the Polak-Ribiere formula (Fletcher 1980, p. 66).

An example

As an initial trial of the algorithm sketched above, we applied the method sketched in the preceeding paragraphs to the estimation of the velocity model depicted in Figure 1, from data computed from the reflectivity in Figure 2. Both parts of the primaries-only acoustic model in this example are layered. Therefore the seismograms are identical functions of time and offset, completely independent of shot location. Evidently, so long as the trial velocities in the inversion are constrained to be layered, otherwise unconstrained reflectivity estimates will be horizontal translates of each other. This suggests that all differential semblance information is already obtained by comparing two reflectivities from neighboring shots. We used a two-shot "line", with shots spaced 100 m apart, and so reduced the computational expense of the trial to a minimum. A data gather is depicted in Figure 3. The data have been muted to remove the direct wave and postcritical energy. The shot depth was 8 m, the receiver depth 12 m, receiver spacing 50 m, near offset 150 m and far offset 1800 m. The sources and receivers were isotropic and punctual (no arrays) with the source time function being a Ricker wavelet with peak frequency 15 Hz. The top surface is pressure-free, and (2, 4) centered finite differences were used in all simulations.

The initial estimate of velocity was constant, $v_1 = 1500m/s$. The velocity estimate after 6 nonlinear conjugate gradient iterations appears in Figure 1 as the dashed curve. Inspection of the estimated reflectivity (Figure 4) reveals that the latter is almost perfectly *flat*: therefore all kinematic detail has been extracted from the data. The fit-to-data is roughly 86%, as indicated in Figures 5 and 6. The slight shallow compression of the inverted reflectivity trace displayed as the dashed line in Figure 2 therefore reveals the limits of kinematic resolution offered by this data set. Higher frequencies, longer times, inclusion of more shot records, and larger offsets might possibly refine this resolution limit.

Conclusion

We have devised an implementation of shot-gather 2D acoustic inversion by differential semblance optimization, and tested it successfully for the computationally inexpensive recovery of a laterally homogeneous model. In this

special case at least the behoviour of the algorithm conformed entirely to the theoretical prediction: all kinematic information content was extracted in a small number of local optimization steps. Other theoretical work and numerical tests have suggested that the smoothness and convexity properties of the DSO misfit function hold also for laterally heterogeneous models (Symes 1991, 1992), so we expect similar behaviour of the inversion algorithm.

Acknowledgments: This work was partly supported by the National Science Foundation, the Office of Naval Research, the State of Texas, and the sponsors of The Rice Inversion Project: Amoco Research, Conoco Inc., Mobil Research and Development Corp., Exxon Production Research Co., Earth Modeling Systems, and Cray Research Inc.

References

[1] G. Chavent and C. Jacewitz 1990, Automatic determination of background velocities by multiple migration fitting, extended abstract, 60th Annual Meeting, Society of Exploration Geophysicists, 1263-1266.

[2] R. Fletcher 1980, *Practical methods of optimization, I: unconstrained optimization*, Wiley, New York.

[3] M. Kern and W. W. Symes 1992, Implementation of differential semblance optimization for 2D velocity estimation, Technical Report 92-3, Department of Mathematical Sciences, Rice University, Houston TX.

[4] P. Kolb, F. Collino, and P. Lailly 1986, Prestack inversion of a 1D medium, *Proc. IEEE 74*, 498-506.

[5] P. Lailly 1983, The seismic inverse problem as a sequence of before-stack migrations, in: *Conference on Inverse Scattering: Theory and Application*, ed. J. B. Bednar, R. Redner, E. Robinson, and A. Weglein, SIAM, Phliadelphia, 206-220.

[6] J. A. Scales, M. A. Smith, and T. L. Fischer 1991, Global optimization methods for highly nonlinear inverse problems, in: *Mathematical and numerical aspects of wave propagation phenomena*, ed. G. Cohen, L. Halpern, and P. Joly, SIAM, Philadelphia, 434-444.

[7] P. L. Stoffa and M. K. Sen 1991, Nonlinear multiparameter optimization using genetic algorithms: inversion of plane-wave seismograms, *Geophysics 56*, 1794-1810.

[8] W. W. Symes 1991, The reflection inverse problem for acoustic waves, in: *Mathematical and numerical aspects of wave propagation phenomena*, ed. G. Cohen, L. Halpern, and P. Joly, SIAM, Philadelphia, 423-433.

[9] W. W. Symes 1992, A differential semblance criterion for inversion of multioffset seismic reflection data, *J. Geophys. Res.*, in press.

[10] W. W. Symes and J. J. Carazzone 1991, Velocity inversion by differential semblance optimization, *Geophysics 56*, 654-663.

[11] A. Tarantola, 1986, A strategy for nonlinear elastic inversion of seismic reflection data, *Geophysics 51*, 1893-1903.

[12] A. Tarantola, 1991, Monte Carlo analysis of geophysical inverse problems, expanded abstract, 61st Annual Meeting, Society of Exploration Geophysicists, 640.

[13] R. Versteeg and G. Grau 1991, *Practical aspects of inversion: the Marmousi experience*, European Association of Exploration Geophysicists, The Hague.

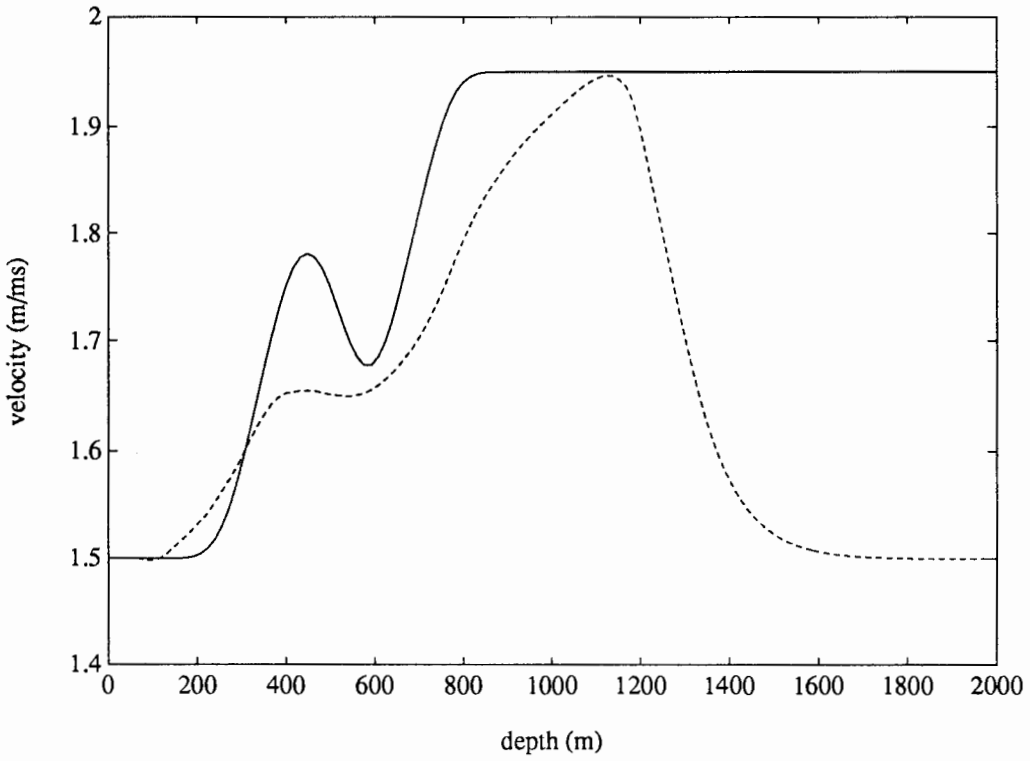

Figure 1: Solid line: target velocity profile. Dashed line: estimated velocity profile after six nonlinear conjugate gradient steps. Note that no moveout information is available to contrain the velocity below roughly 1300 m.

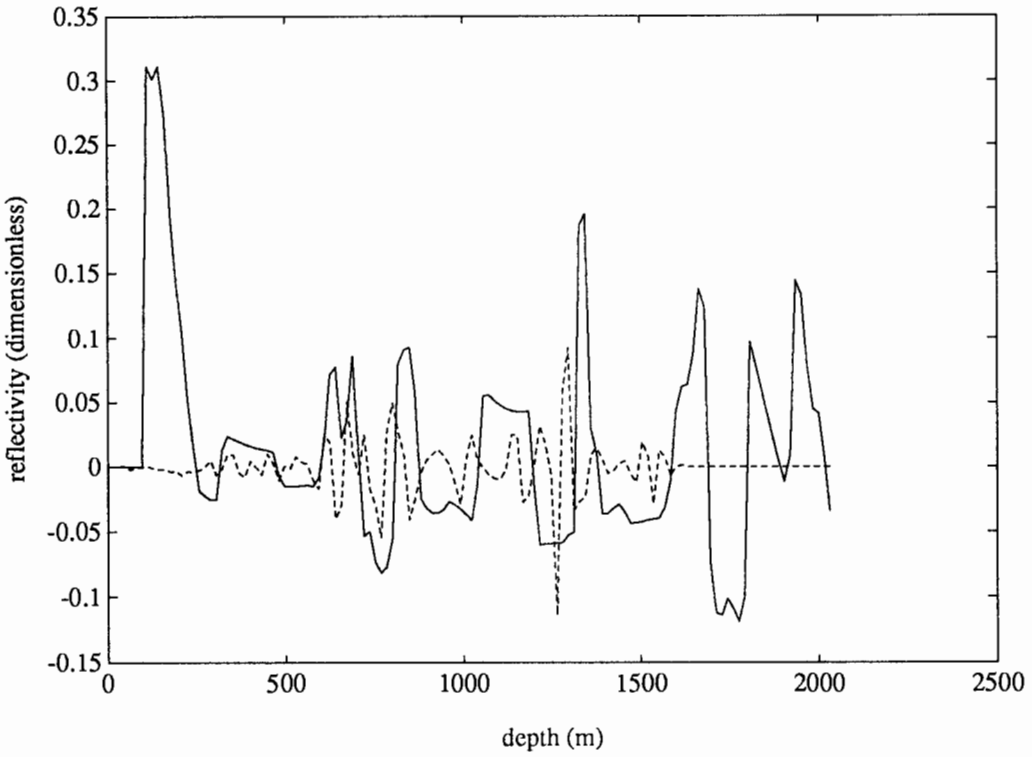

Figure 2: Solid line: target reflectivity profile. Dashed line: trace of estimated reflectivity profile after six nonlinear conjugate gradient steps. Trace shown is located at 3584 m from the west edge of the model (roughly the center of Figure 4). The passband-filtering effect of the inversion is evident, as is a compression of events toward the surface, amounting to roughly 50 m depth error at 1300 m. At this surface location the inversion aperture does not produce images of shallower reflectors (see Figure 4).

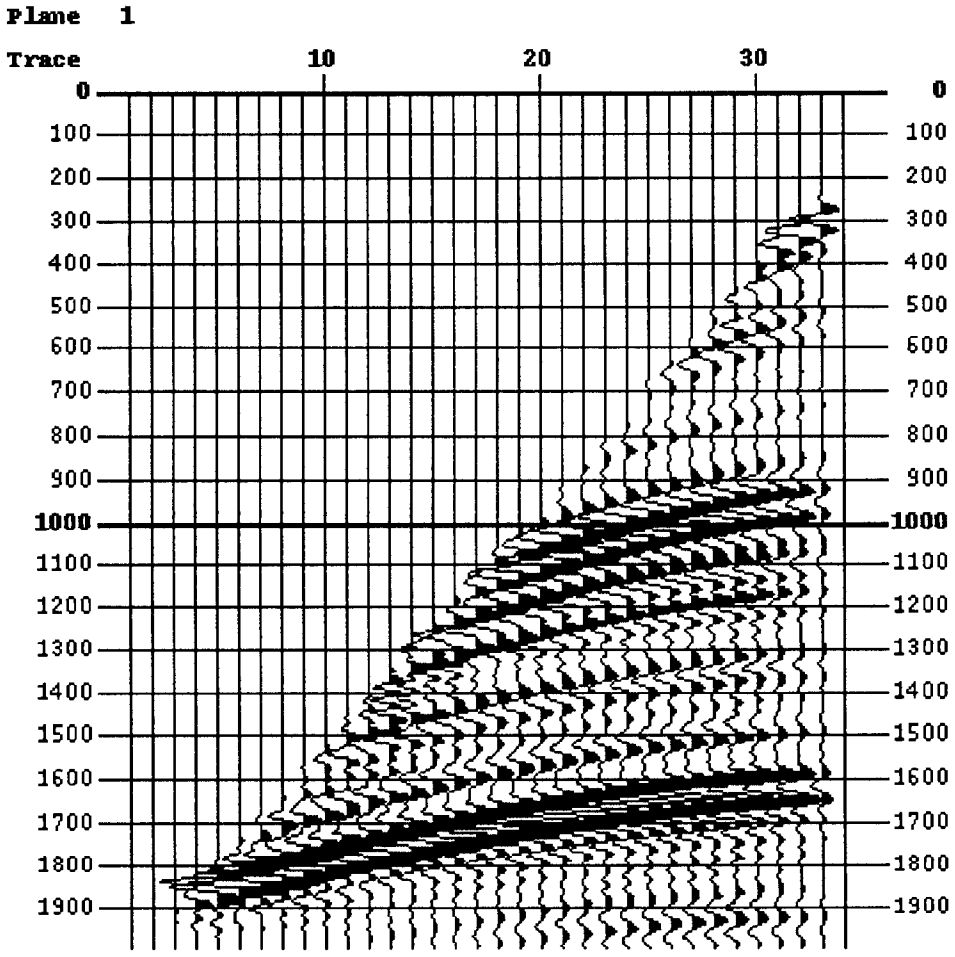

Figure 3: Data gather for shot at 3000 m from left edge of model. Near offset 150 m, far offset 1800 m, receiver spacing 50 m. Source is 15 Hz zero-phase Ricker wavelet. This is the first of two shot gathers in the "line". The second is identical, and is positioned at 3100 m.

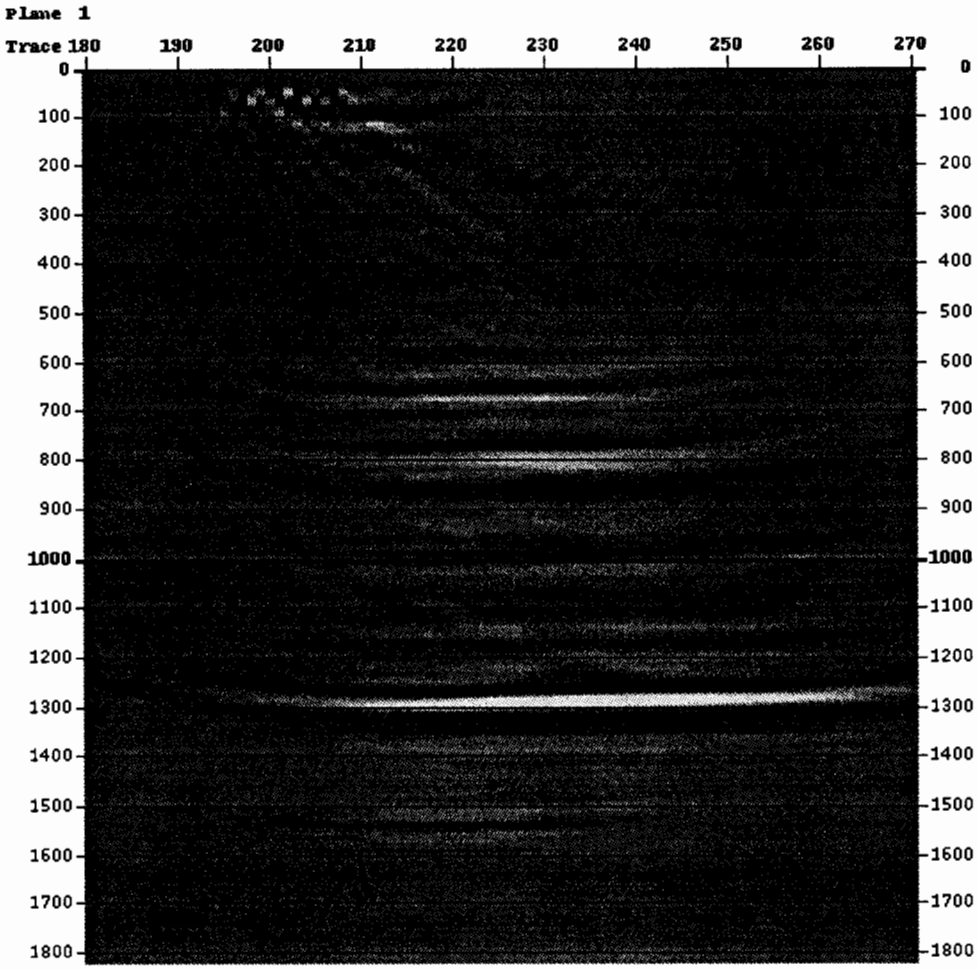

Figure 4: Reflectivity estimated after six steps of nonlinear conjugate gradient iteration. The events are essentially flat, which suggests the kinematical correctness of the inverted velocity (Figure 1). Trace spacing is 16 m; vertical axis is in meters.

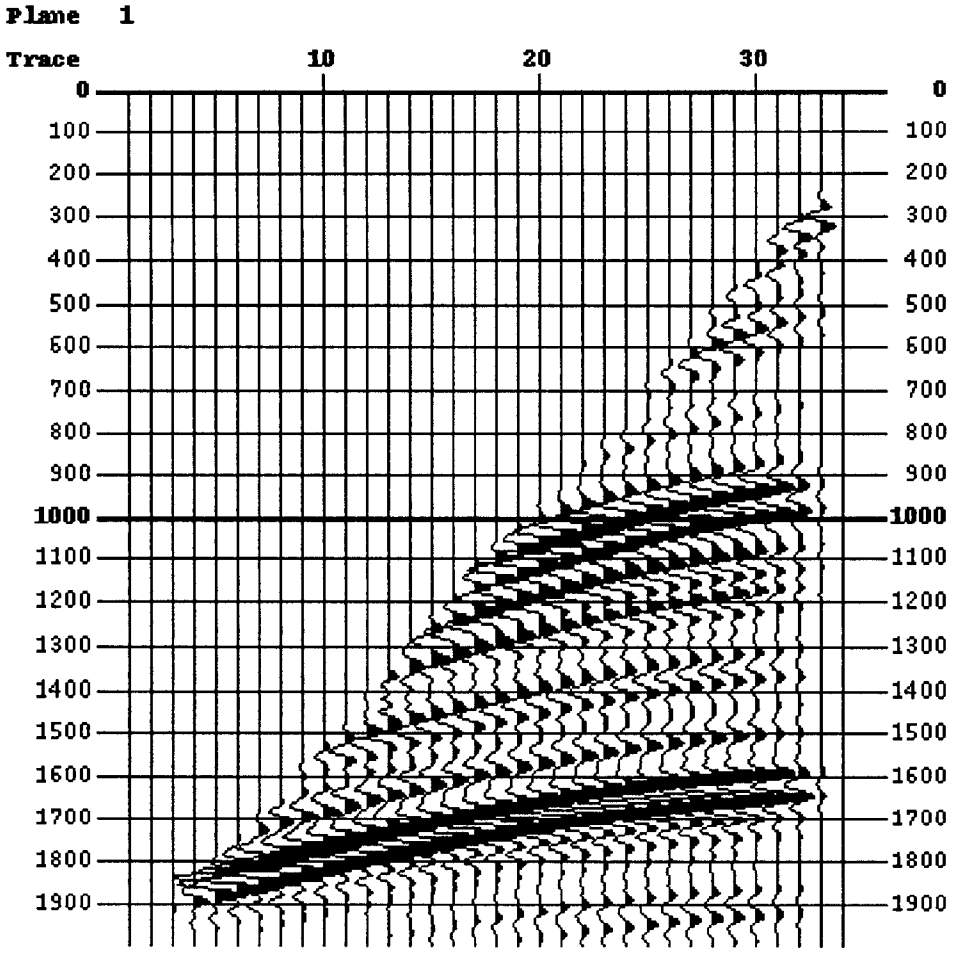

Figure 5: Predicted data gather for shot at 3000 m from left edge of model. Compare Figure 3.

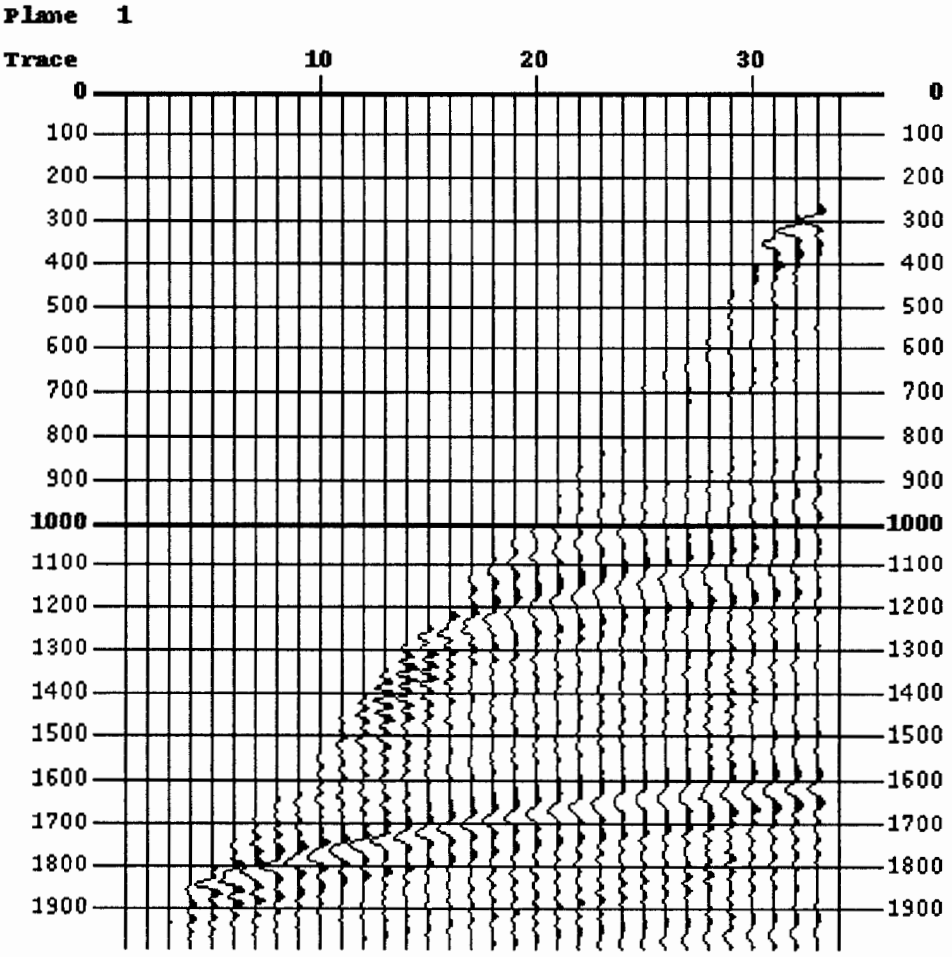

Figure 6: Difference between data and predicted shot gathers. RMS error is 14%.