GEOPHYSICS, VOL. 54, NO. 8 (AUGUST, 1989); P. 984–990, 13 FIGS., 1 TABLE.

Reference velocity model estimation from prestack waveforms: Coherency optimization by simulated annealing

Evgeny Landa*, Wafik Beydoun‡, and Albert Tarantola§

ABSTRACT

Coherency inversion, which consists of maximizing a semblance function calculated from unstacked seismic waveforms, has the potential of estimating reliable velocity information without requiring traveltime picking on unstacked data.

In this work, coherency inversion is based on the assumption that reflectors' zero-offset times are known and that the velocity in each layer may vary laterally. The method uses a type of Monte Carlo technique termed the generalized simulated annealing method for updating the velocity field. At each Monte Carlo step, time-to-depth conversion is performed. Although this procedure is slow at convergence to the global minimum, it is robust and does not depend on the initial model or topography of the objective function. Applications to both synthetic and field data demonstrate the efficiency of coherency inversion for estimating both lateral velocity variations and interface depth positions.

INTRODUCTION

Good knowledge of the velocity field is necessary for obtaining accurate images of an earth model with migration and linearized inversion techniques. This information can be obtained by different kinematic methods, such as traveltime tomography (Bishop et al., 1985; Chiu et al., 1986). The combination of traveltime tomography and depth migration has also been used to produce enhanced subsurface images (Bording et al., 1987; Stork and Clayton, 1985; Stork, 1988). However, in practice, traveltime tomography suffers from a serious shortcoming: its strong requirement for accurate prestack reflection traveltime picking. An interesting approach is described by van der Made (1988), in which stacking velocities and zero-offset times are used to construct a subsurface macro-model. This technique is based essentially on a hyperbolic approximation to arrival times.

To overcome the need for accurate traveltime picking, Landa et al. (1988) have proposed a coherency inversion method which does not depend on prestack time picking and is not based on curve fitting or hyperbolic approximations. The method is formulated as an optimization algorithm producing a velocity model which maximizes some measure of coherency (semblance function). This measure is calculated on unstacked trace gathers (common-shotpoint or common-midpoint) in a time window along the traveltime curves computed by tracing rays through the model. Positions of interfaces and interval velocities within the layers are represented by spline functions defined by their node points. Unknown parameters are values at nodes of velocity and interface depth position. Semblance optimization has an advantage over traveltime interpretation for low signal-to-noise ratio data, when prestack traveltime picking is an unreliable process.

In this paper, we consider the special case where zero-offset traveltimes are known for principal reflectors (for example, from poststack picking). This knowledge allows a great simplification of the optimization process, because we can alternatively update velocities (using the coherency measure) and interface position (using zero-offset time information) until the optimal solution is obtained.

Since velocity updating is a highly nonlinear process, it is done by a sophisticated version of a Monte Carlo technique, namely, generalized simulated annealing. The method is a random walk that samples the objective function in the space of independent variables. The advantage of this method is the ability to migrate through a sequence of local extrema and to recognize when the global extremum has been reached. Simulating annealing was introduced by Kirkpatrick et al. (1983) and was first used in geophysics by Rothman (1985). A review of this method can be found in Tarantola (1987).

Manuscript received by the Editor July 18, 1988; revised manuscript received January 30, 1989.
*Institut de Physique du Globe de Paris, 4 place Jussieu, F-75252 Paris Cedex 05; also Elf Aquitaine, 64000 Pau, France.
‡Elf Aquitaine, 64000 Pau, France; Earth Resources Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02142; ARCO Oil and Gas Company, 2300 West Plano Parkway, Plano, TX 75075.
§Institute de Physique du Globe de Paris, 4 place Jussieu, F-75252 Paris Cedex 05, France.

We discuss briefly the limitations of the proposed method and demonstrate the application of velocity inversion to field data.

PROCEDURE

Coherency inversion (Landa et al., 1988) can be described as follows:

The real medium is modeled as a series of layers separated by interfaces. Velocity within each layer can vary laterally. Interfaces and velocities are represented by spline functions. Unknown parameters in inversion are vertical node locations for interfaces (vector **Z**) and values of velocities at nodes (vector **V**). The interpretive model is now characterized by the vector of parameters $\mathbf{m} = \{\mathbf{V}, \mathbf{Z}\}$.

Inversion, performed iteratively layer after layer, consists of finding a velocity-depth model which maximizes semblance functional $E(\mathbf{m})$ calculated for all prestack trace gathers in a time window along traveltime trajectories generated by ray tracing. This optimization process can be defined as the maximization of the function $E(\mathbf{m})$ given by

$$E(\mathbf{V}, \mathbf{Z}) = \sum_i \sum_{k=0}^{K} \frac{\left\{ \sum_j U_{ij}[k\delta t + \tau(\mathbf{V}, \mathbf{Z})] \right\}^2}{\sum_j \left\{ U_{ij}[k\delta t + \tau(\mathbf{V}, \mathbf{Z})] \right\}^2}, \qquad (1)$$

where U_{ij} represents the seismic trace for the ith source and jth receiver, τ is the traveltime calculated by ray tracing through the model, δt is the sample interval, and $K\delta t$ is the time window for semblance calculation.

A possible difficulty in the practical application of coherency inversion is the inclusion in the semblance of multiple reflections and noise. To prevent this, coherency inversion should be used interactively so that a skilled interpreter can avoid including noisy data or multiples.

For traveltime calculation, we applied a ray-tracing algorithm described in Haas et al. (1987) based on a ray-continuation technique. The maximization of expression (1) in order to obtain the best model $\mathbf{m} = \{\mathbf{V}, \mathbf{Z}\}$ could be performed by using any optimization method. In fact, we show below that the function $E(\mathbf{m})$ has secondary maxima. Here we choose to simplify the optimization process by using, in addition to the prestack seismic records, the complementary data set of zero-offset traveltimes $\mathbf{T}_0(x)$ for selected reflectors; x represents horizontal position. $\mathbf{T}_0(x)$ can be extracted or picked on stacked time sections. Conventional processing, such as muting, filtering, static corrections, deconvolution, multiple attenuation, etc., can be applied to prestack seismic data.

Our idea is to use this time information for a rough estimate of the interface depth position, while using the prestack seismic data for estimating the velocity functions. It turns out that the first subproblem is quite easy to solve using gradient optimization methods, while the second subproblem is highly nonlinear and requires the use of a Monte Carlo technique of global optimization. Depth conversion and velocity inversion are obviously not completely independent, however. The iterative use of these two processes leads, in general, to efficient velocity estimation.

The optimization strategy we use is depicted in Figure 1. Assume we have a velocity model **V**; at the very beginning, this is simply a rough model, for instance, horizontal layers with constant velocities. First, we convert the $\mathbf{T}_0(x)$ information into $\mathbf{Z}(x)$ information. This can be done using the well-known ray migration algorithm, but we prefer to use the procedure (Keydar et al., 1989) of minimizing the objective function

$$S(\mathbf{Z}) = \sum_i \left[\mathbf{T}_0^{\text{obs}}(x_i) - \mathbf{T}_0^{\text{cal}}(x_i, \mathbf{Z}) \right]^2. \qquad (2)$$

Equation (2) means that, given a velocity model **V**, we can compute a depth position model $\mathbf{Z}(\mathbf{V})$, using the complementary $\mathbf{T}_0(x)$ information. If **Z** is a function of **V**, the semblance functional (1) depends only on the variable **V** and turns out to be a highly nonlinear functional with secondary maxima.

One of the main advantages of separating structural and velocity parameter estimations can be observed in the presence of complex structural elements, such as pinchouts or faults. The location of such elements is estimated by $\mathbf{T}_0$ time-to-depth conversion [expression (2)] using constrained optimization (Keydar et al., 1989). We can also use a depth migration algorithm for an imaging step (Stork and Clayton, 1985; Bording et al., 1987). In this case, interface depth positions (including pinchouts and faults) are picked directly from the depth section by the interpreter.

The key question in realizing the strategy described above is "How do we update velocity vector **V** in order to reach the global maximum of the objective function E?" To answer this question, we calculated expression (1) for a simple model with one layer and a velocity function described by a spline with only two nodes. Figure 2 shows a contoured map for the velocities at the two nodes corresponding to ten

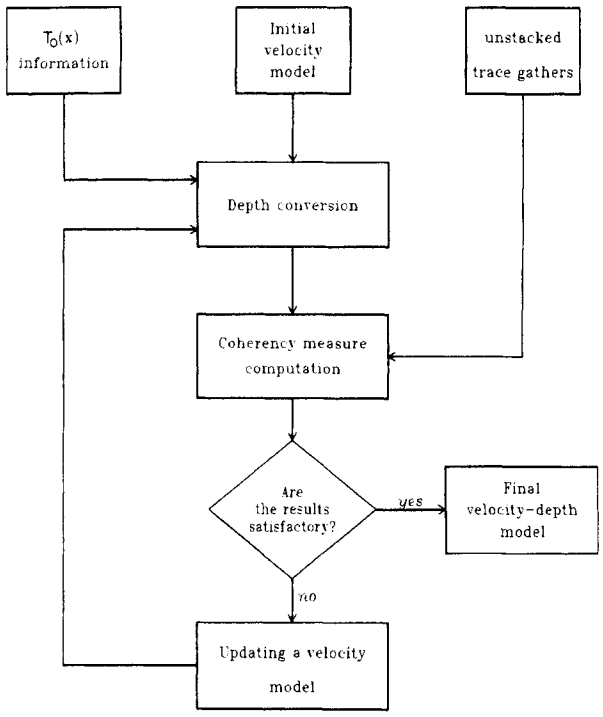

FIG. 1. Block scheme of reference model definition.

common-midpoint (CMP) gathers. The true values of velocity at the nodes are 3000 m/s and 3200 m/s. Note that even in this very simple case the objective function has a "valley" structure and (what is more important) is multimodal. These circumstances make it practically impossible to use most algorithms of nonlinear programming for maximizing the coherency function (1).

In this case, the most appropriate solution is obtained using Monte Carlo-type techniques. We use the simulated annealing method as described by Rothman (1985). Simulated annealing is a convenient way to find a global extremum of a function that has several local extrema and may not be smooth. Simulated annealing is a type of Monte Carlo (or random search) technique that simulates the process by which a crystal is grown from a melt. The crystallization process is analogous to optimization in that both processes tend from a disordered to an ordered system. The probability distribution used in simulated annealing is the damped exponential function used in statistical mechanics. The method is implemented in the following manner. Starting with initial values of unknown variables, the corresponding initial value of the objective function E_0 is calculated. Then a random point is chosen on the surface of the n-dimensional hypersphere, where n is the dimension of the problem. A procedure for obtaining a random direction is to choose n random numbers from the uniform distribution and convert them into direction cosines. In this direction a step of size s is taken. The new value of the objective function E_1 is calculated. The step is accepted with probability p given by

$$p = 1 \qquad \text{if } \Delta E = E_1 - E_0 \geq 0$$

$$p = \exp(\beta \Delta E) \qquad \text{if } \Delta E < 0,$$

where β is a positive parameter. This means that the steps when $E_1 > E_0$ are accepted unconditionally ("standard" Monte Carlo method). Unsuccessful steps ($E_1 < E_0$) are accepted according to the following additional experiment: A random number ρ is generated from the uniform distribution on (0, 1) and is compared to the value $\exp(\beta \Delta E)$. If $\rho < \exp(\beta \Delta E)$, then the step is accepted; otherwise it is rejected. The probability of accepting an unsuccessful step is always greater than zero, so the path may walk out of local minima. The number of steps required for that depends on the value of β. This parameter must be adjusted for each particular objective function and is obtained after some experimentation (Rothman, 1985). (In the physics of annealing, $\beta = 1/kT$ where k = Boltzmann's constant and T = temperature.)

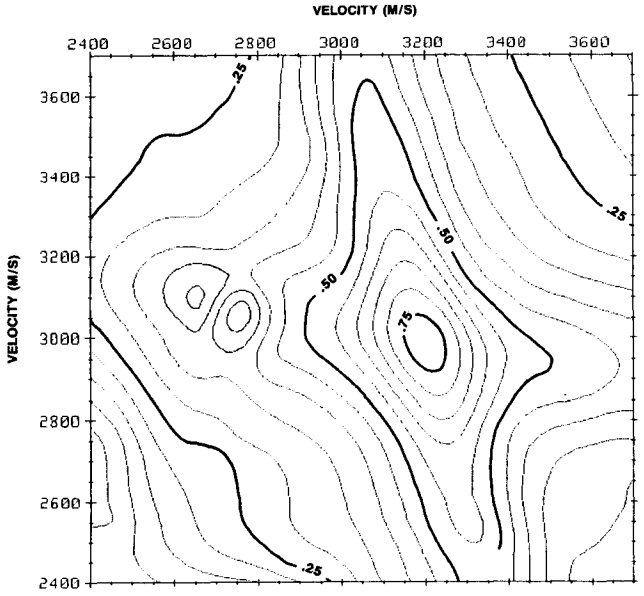

FIG. 2. Contours of the objective function (1) for the velocities of the two nodes.

For practical reasons, we use the generalized simulated annealing method (Bohachevsky et al., 1986). In this method, the probability of accepting a detrimental step tends to zero as a random walk approaches the global extremum. Let us assume that E_{max} is a known value of the objective function at the global maximum. The required behavior of the probability is easily accomplished by setting $p = \exp[\beta(E_{max} - E)^q \Delta E]$, for $\Delta E < 0$, where q is an arbitrary negative number; for $q = 0$, we have the standard simulated annealing method. We can summarize the generalized simulated annealing method in the following manner:

Assume $E(\mathbf{m})$ is the function to be maximized.

(1) Let $\mathbf{m}_0$ be an arbitrary starting point.
(2) Set $E_0 = E(\mathbf{m}_0)$. If $|E_0 - E_m| < \varepsilon$, stop.
(3) Determine random direction $\mathbf{r}$.
(4) Set $\mathbf{m}_1 = \mathbf{m}_0 + \delta\mathbf{m} \cdot \mathbf{r}$, where $\delta\mathbf{m}$ is a step.
(5) Calculate $E_1 = E(\mathbf{m}_1)$. If $E_1 > E_0$, set $\mathbf{m}_0 = \mathbf{m}_1$ and $E_0 = E_1$. If $|E_0 - E_{max}| < \varepsilon$, stop. Otherwise, go to step (3).
(6) If $E_1 < E_0$, set $p = \exp[\beta(E_{max} - E_1)^q \Delta E]$. Generate a uniform (0–1) variable ρ. If $\rho > p$, go to step (3). If $\rho < p$, set $\mathbf{m}_0 = \mathbf{m}_1$, $E_0 = E_1$, and go to step (3).

In practice, the random walk is terminated when an arbitrary number of trials (20–50) fails to produce an acceptable point.

EXAMPLES

Synthetic example I

We applied our inversion technique to the three-layered model shown in Figure 3. Interval velocities in the first two layers are constant and equal to 1500 m/s and 2000 m/s, respectively. Lateral changes of velocity in the third layer are described by a spline function with six nodes at horizontal locations: 2000, 3000, 4000, 5000, 6000, and 7000 m; the exact values of velocity are 3000, 3500, 3000, 3500, 3000, and 3500 m/s, respectively (dashed line in Figure 3). For this model, we generated 21 CMP gathers by putting the Ricker wavelet of unit amplitude at the positions corresponding to the reflection traveltimes calculated using the ray-tracing program and adding random noise. The first CMP is located at 2000 m and the last at 7000 m. Each gather comprises 24 traces with a trace interval of 100 m and a minimum offset of 0 m. Our targets in this example were velocity values and reflector positions at the node points for the third layer. Input data consisted of CMP gathers and correct zero-offset times. To this data set we applied our optimization scheme

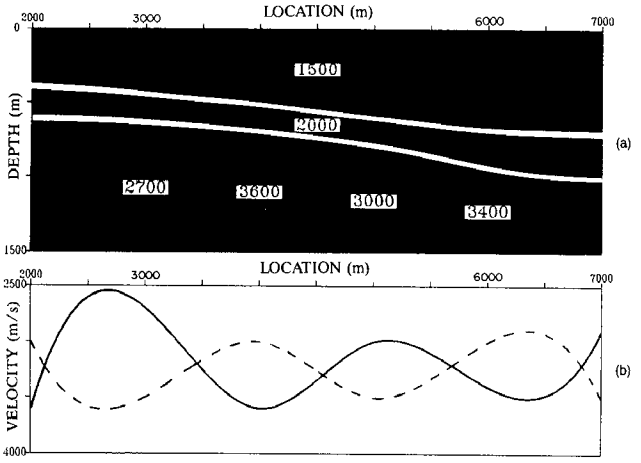

FIG. 3. (a) Exact model for synthetic example I. (b) The exact (dashed line) and initial (solid line) velocity for the bottom layer.

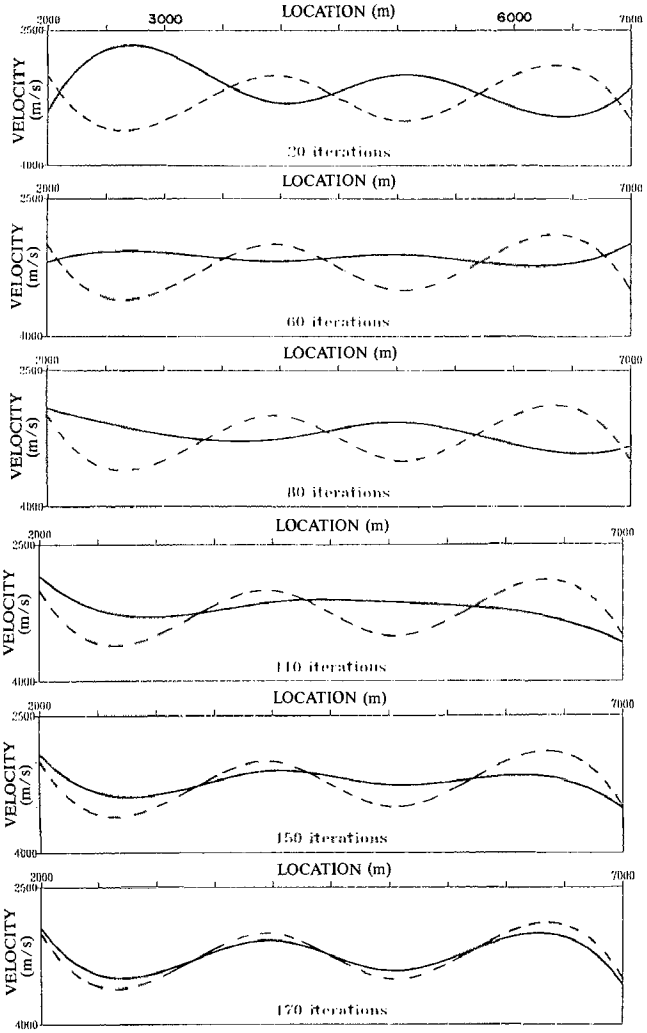

FIG. 4. Convergence of the inversion process by the generalized simulated annealing method (true velocity is shown as dashed lines), for synthetic example I.

with the initial velocity model shown in Figure 3 (solid line). After several trials, the following parameters were chosen for the simulated annealing method: step $s = 200$ m/s, $q = -2$, and $\beta = 7$. Figure 4 illustrates convergence of the inversion process after 20, 60, 80, 110, 150, and 170 iterations. We terminated the process when the semblance exceeded the threshold value of 0.65. The final velocity function is very close to the true one (dashed line in Figure 4). Each iteration in this case consists of about 2–5 normal-incidence ray-tracing steps and one CMP ray-tracing step. To illustrate the efficiency of the generalized simulated annealing optimization method, we performed the same inversion example using the standard Monte Carlo for maximizing the coherency functional. In "standard Monte Carlo method," we just set to zero the probability of acceptance of unsuccessful steps. We terminated the process after 260 iterations and the results are shown in Figure 5. Figure 6 illustrates the comparison of these two methods. After about 80 iterations there is no change in the velocity function, which indicates the existence of a local minimum and the inability of the standard Monte Carlo method to escape from it. Of course, for a large enough step s, it could escape from local minima; but then the number of iterations would increase significantly.

Synthetic example II

The exact velocity-depth model for this example is shown in Figure 7 (dashed line). As in the previous example, the first two layers are homogeneous and the lateral variation in the third layer is defined by a spline function with five nodes. For this model, 40 common-shot gathers were computed using the elastic paraxial ray method (Beydoun and Keho, 1987). Each shot gather comprises 12 traces with 150 m spacing. We sorted these data into 21 CMP gathers with six traces in each. The first CMP is located at 2000 m and the last CMP is at 7000 m. The distance between traces in the CMP gathers is 300 m, the minimum offset being 0 m. We added a small amount of random noise to this data set. Zero-offset traveltimes for the third reflector taken from the forward modeling program output were also used as input for the inversion. The initial velocity model for the third layer was obtained by coherency inversion assuming a constant-velocity model and was equal to 2900 m/s, representing the average velocity in the layer (solid line in Figure 7). To estimate the lateral variations of the velocity function, the inversion scheme was applied to the data set, where the parameters for generalized simulated annealing were taken as in the previous example. The process was automatically terminated after 150 iterations due to inability to find a new acceptable point during the last 40 iterations. The resulting velocity-depth model is shown in Figure 8. Our final result describes satisfactorily the behavior of the velocity at almost all points except at the left and right boundaries of the model. Differences between true velocity (dotted line in Figure 8) and the result of inversion (solid line in the figure) at the left part of the model can be explained by edge effects, since we do not have sufficient data to correct velocity estimates. To understand the nature of differences in the right part of the model, Figure 9 displays a CMP gather located in this area. The traveltime curve and the time window used for the semblance computation for the last iteration are shown by the thick line.

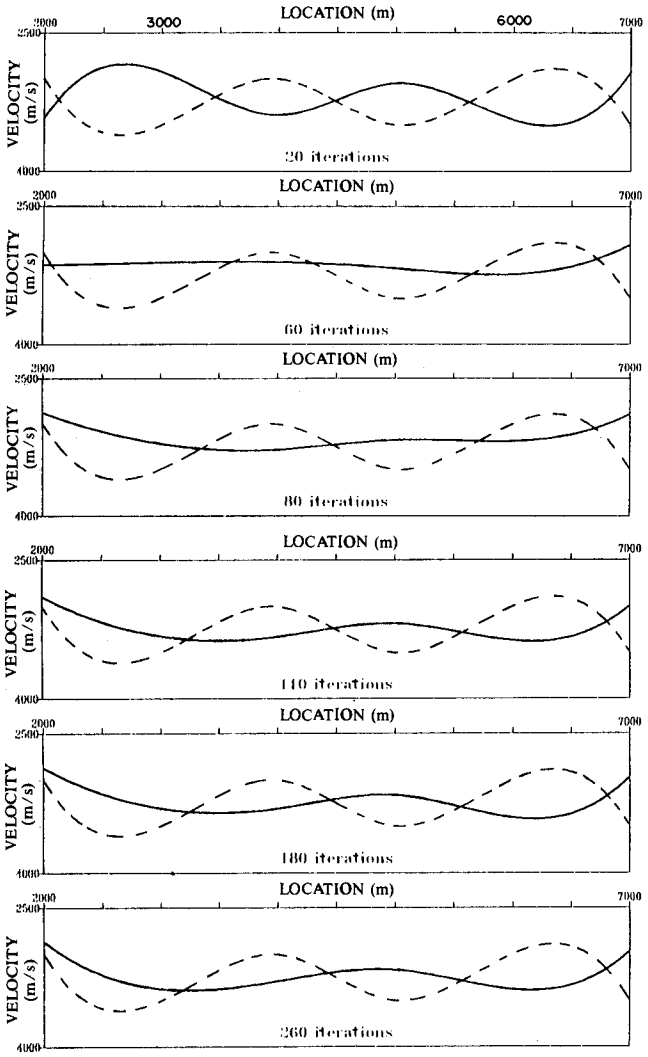

FIG. 5. Inversion of synthetic example I using a Monte Carlo method. The velocity does not change after 80 iterations (true velocity is shown as dashed lines).

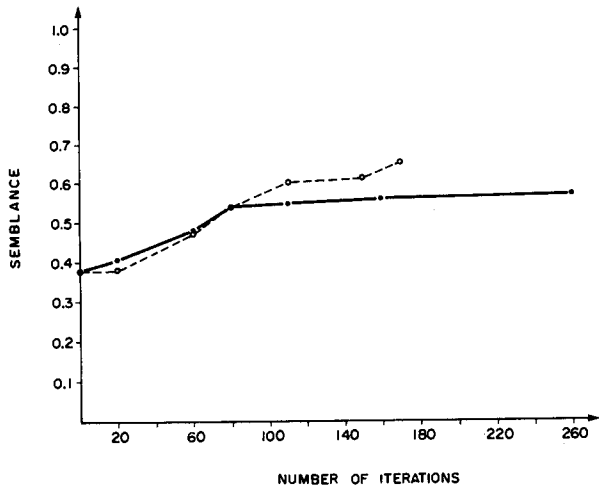

FIG. 6. Comparison for synthetic example I between the convergences of the generalized simulated annealing method (dashed line) and a standard Monte Carlo method (solid line).

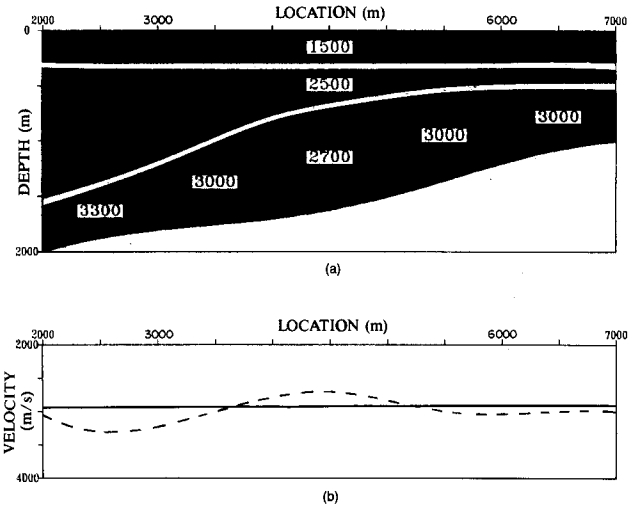

FIG. 7. Velocity-depth model for synthetic example II. (a) Actual model, (b) exact (dashed) velocity in the third layer. The solid line is a constant-velocity inversion for layer 3.

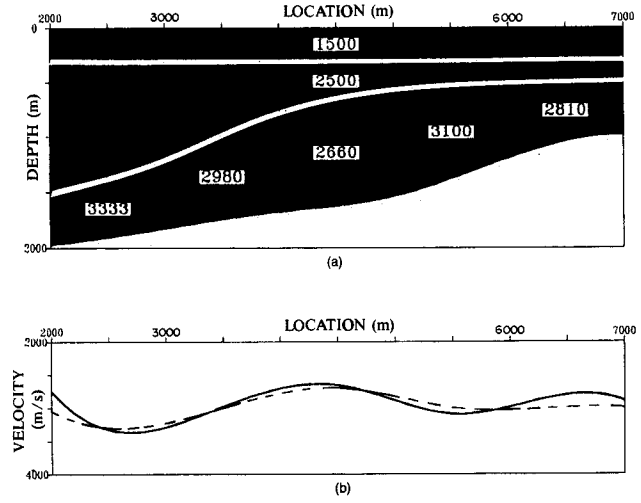

FIG. 8. Final velocity-depth model for synthetic example II after 150 iterations. Same convention as in Figure 7.

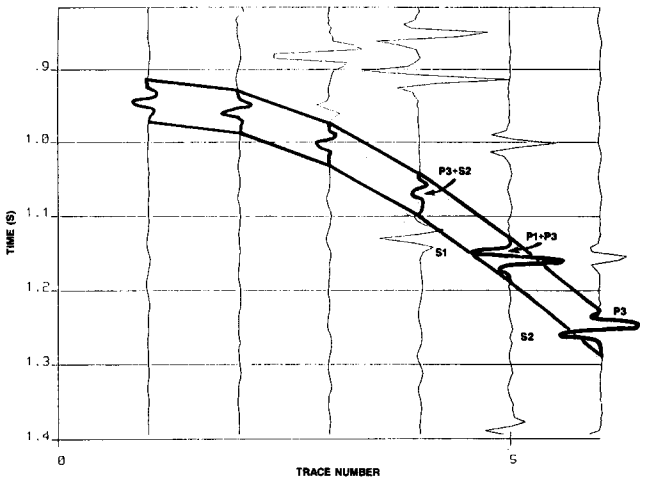

FIG. 9. CMP gather for synthetic example II at CMP 6600. The time window for the semblance computation is shown in thick lines.

We can see that (1) there is interference between events on the last traces of this gather and that (2) there is a phase change of the reflected wave from the third interface ($\mathbf{P}_3$) at the last two traces, due to the supercritical angles of reflection. These two effects lead to incorrect velocity estimation for the fifth node point of the spline function. For this case, we calculated error values in velocity estimates at each of the five node points using the response surface technique of Bard (1974). This technique consists of estimating the diagonal elements of the objective function's Hessian matrix. Results are presented in Table 1. The velocity estimation error for the point with x-coordinate 6500 m is about three times larger than for other points.

A more appropriate way of handling the amplitude changes would be to consider an objective function evaluated from differences between observed and calculated amplitudes. Seismograms, or preferably envelopes, would then be calculated and compared to data, leading to an objective function of the type used by Tarantola (1987).

Table 1. Results of inversion and error estimates for the third layer of synthetic example II.

$E = 0.58$		
X coordinate (m)	Velocity (m/s)	Errors (m/s)
2000	2761	129
2500	3333	133
3500	2980	99
4500	2657	140
5500	3100	123
6500	2808	464
7500	2905	117

Field data example

Following the successful experiments on synthetic data, we applied 2-D velocity inversion to real data. The stacked section of a land seismic line is shown in Figure 10. Five reflection events were picked by interpreting this section and using a priori geologic information. At the right side of the section, events have zero-offset times of about 0.75, 1.3, 1.8, 2.3, and 2.5 s, respectively. Our target in this example was estimation of lateral velocity variations within layers 3, 4,

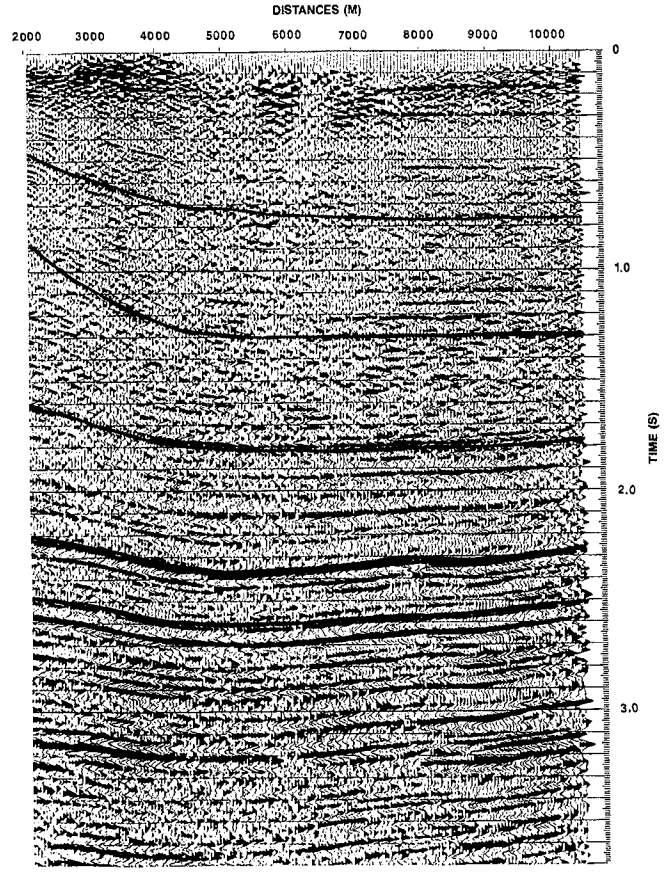

FIG. 10. Stacked time section for real data example.

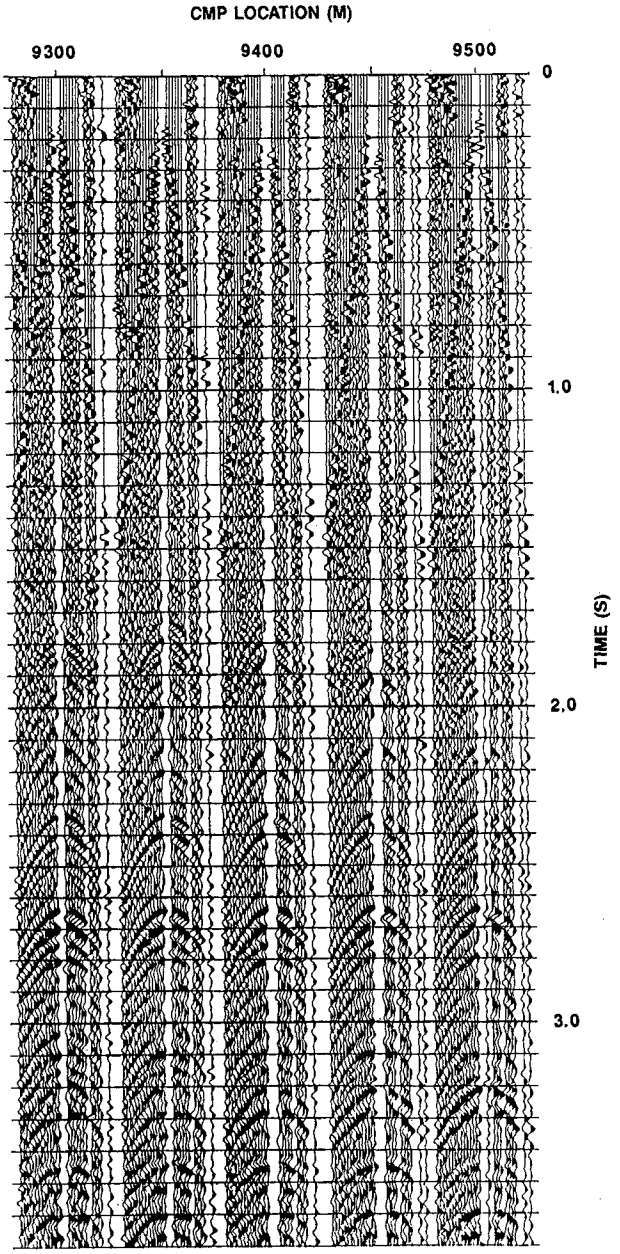

FIG. 11. Five CMP gathers for real data example.

and 5. The input data consist of 25 CMP gathers, located between 2000 and 10 000 m, picked zero-offset times, and the initial model. Five CMP gathers are shown in Figure 11. For inversion, from each gather we used 15 traces with 150 m spacing and minimum offset equal to 250 m. The first two layers were assumed to be known from conventional processing and did not participate in the inversion. The initial model for the third, fourth, and fifth layers was chosen to be stratified with constant velocities equal to 3700 m/s, 3100 m/s, and 3800 m/s, respectively. Inversion was performed layer after layer. Unknown parameters (spline node points) for velocities were chosen: five for the third layer and four for the fourth and fifth layers. Nodes for the depth function were taken every 2000 m. The final velocity-depth model is shown in Figure 12. Velocity variations for the fifth layer are very small and are not shown in the figure. Figure 13 illustrates two time windows on the CMP gather at 9800 m used for coherency calculation for the fourth and fifth reflectors at the last iteration. The calculated traveltime curves clearly follow very good reflection events.

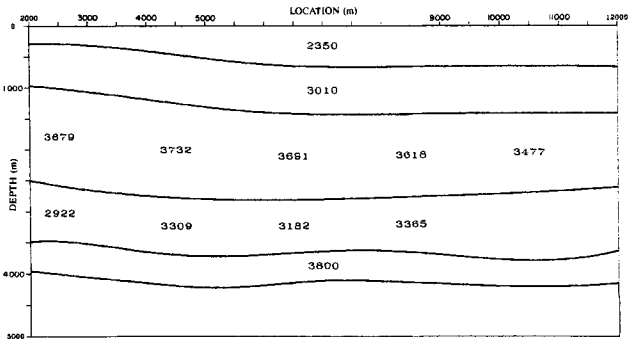

FIG. 12. Final velocity-depth model for real data example.

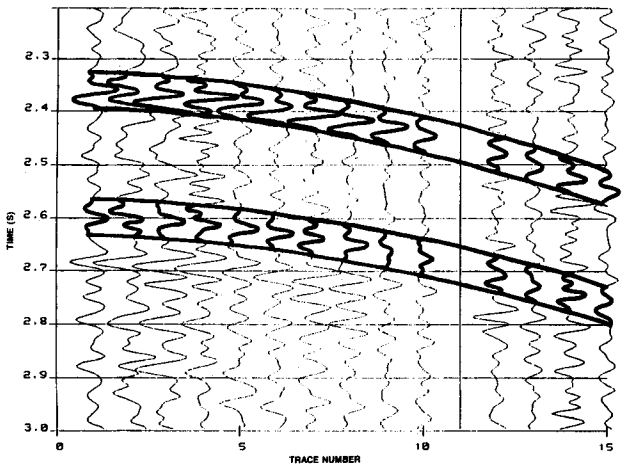

FIG. 13. CMP gather at 9800 m for real data example. Two time windows for the semblance computation including the reflections from the fourth and fifth interfaces are shown in thick lines.

CONCLUSIONS

A method of 2-D velocity inversion or reference-model estimation by coherency optimization has been presented. It is based on separating structural and velocity parameter estimation. A generalized simulated annealing method is used to maximize the semblance function. Although generalized simulated annealing slowly converges to the global optimum, it is practically independent of the initial model. The inversion was successfully applied to synthetic and real data.

A natural extension of this technique would be to take waveforms or, possibly, their envelopes into account. Calculation of amplitudes will then be necessary, leading to a seismogram misfit objective function of the type used by Tarantola (1987).

ACKNOWLEDGMENTS

This research was supported by Elf Aquitaine and by sponsors of the Equipe de Tomographie Geophysique at the Institute de Physique du Globe de Paris (Amoco, Aramco, CGG, Convex, Elf Aquitaine, IFP, Mobil, Shell, Sun Microsystems, and TOTAL). We wish to thank Manuela Mendes for preparing the data of example II using the paraxial ray computer program PRX87, Andre Haas for providing his ray-tracing Fortran code, and Bernadette Lajus for adapting the ray-tracing program of Haas and Viallix (1988) to the coherency inversion.

REFERENCES

Bard, J., 1974, Nonlinear parameter estimation: Academic Press, Inc.

Beydoun, W., and Keho, T., 1987, The paraxial ray method: Geophysics, **52**, 1639–1654.

Bishop, T., Bube, K., Cutler, R., Langan, R., Love, P., Resnick, J., Shuey, R., Spinder, D., and Wyld, H., 1985, Tomographic determination of velocity and depth in laterally varying media: Geophysics, **50**, 903–923.

Bohachevsky, I., Johnson, M., and Stein, M., 1986, Generalized simulated annealing for function optimization: Technometrics, **28**, 209–217.

Bording, R. P., Gersztenkorn, A., Lines, L. R., Scales, J. A., and Treitel, S., 1987, Applications of seismic traveltime tomography: Geophys. J. Roy. Astr. Soc., **90**, 286–303.

Chiu, S., Kanasewich, E., and Padke, S., 1986, Three-dimensional determination of structure and velocity by seismic tomography: Geophysics, **51**, 1559–1571.

Haas, A., Dezard, Y., Guiziou, T. and Duval, T., 1987, Seismic tomography in heterogeneous media: the stochastic inversion by ray continuation method: 57th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 836–838.

Keydar, S., Koren, Z., Kosloff, D., and Landa, E., 1989, Optimal time-to-depth conversion: Geophysics, **54**, 1001–1005.

Kirkpatrick, S., Gelatt, C., and Vehhi, M., 1983, Optimization by simulated annealing: Science, **220**, 671–680.

Landa, E., Kosloff, D., Keydar, S., Koren, Z., and Reshef, M., 1988, A method for determination of velocity and depth from seismic reflection data: Geophys. Prosp., **36**, 223–243.

Rothman, D., 1985, Nonlinear inversion, statistical mechanics, and residual statics estimation: Geophysics, **50**, 2784–2796.

Stork, C., 1988, Ray tomography method for inversion of reflection seismic data: Ph.D. thesis, Cal Tech.

Stork, C., and Clayton, R. W., 1985, Iterative tomographic and migration reconstruction of seismic images: 55th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 610–613.

Tarantola, A., 1987, Inverse problem theory: Elsevier Science Publ. Co.

van der Made, P., 1988, Determination of macro subsurface models by generalized inversion: Ph.D. thesis, Delft Univ.