Geophysical Prospecting **39,** 599–611, 1991

A SIMULATED ANNEALING APPROACH TO SEISMIC MODEL OPTIMIZATION WITH SPARSE PRIOR INFORMATION[1]

KLAUS MOSEGAARD[2] and PETER D. VESTERGAARD[3]

ABSTRACT

MOSEGAARD, K. and VESTERGAARD, P.D. 1991. A simulated annealing approach to seismic model optimization with sparse prior information. *Geophysical Prospecting* **39,** 599–611.

It is well known that seismic inversion based on local model optimization methods, such as iterative use of linear optimization, may fail when prior information is sparse. Where the seismic events corresponding to reflectors of interest remain to be identified, a global optimization technique is required.

We investigate the use of a global, stochastic optimization method, that of simulated annealing, to solve the seismic trace inversion problem, in which the two-way traveltimes and reflection coefficients are to be determined. The simulated annealing method is based on an analogy between the model-algorithm system and a statistical mechanical system. We exploit this analogy to produce improved annealing schedules. It is shown that even in cases of virtually no prior information about two-way traveltimes and reflection coefficients, the method is capable of producing reliable results.

INTRODUCTION

The seismic trace inversion problem can be formulated as a non-linear search for an acoustic impedance function that is stepwise constant in depth, and whose seismic response, modelled by means of the convolutional model, is as close as possible to the measured response in the least-squares sense. Cooke and Schneider (1983) found that for this problem, a traditional, local optimization method fails if not provided

[1] Received August 1989, revision accepted November 1990, last material received January 1991.

[2] Geophysical Institute, University of Copenhagen, Haraldsgade 6, 2200 Copenhagen N, Denmark.

[3] Imperial College, Department of Geology, Royal School of Mines, Prince Consort Road, London SW7 2BP, U.K. Formerly Ødegaard and Danneskiold-Samsøe Aps, Copenhagen.

with an initial guess for which the two-way traveltimes to the layer interfaces are closer to the true values than the side lobes of the corresponding seismic events.

In cases of good well control, detailed prior information about the subsurface model may be available, and a good initial guess may be made. In such cases, traditional trace inversion methods may prove successful. However, in many exploration problems and in the early phases of oil and gasfield developments, the well control is sparse and a good initial guess may not be available. In such cases the trace inversion problem becomes a global model optimization problem, as many secondary minima for the misfit function exist.

Global optimization problems are much more difficult to solve than local optimization problems, since the local geometry of the misfit function surface in the model space does not directly contain information about the direction to follow in a search for the global minimum. We have therefore applied a stochastic optimization technique, that of simulated annealing, in a global search for the optimal subsurface model. The usefulness of this method, when applied to global optimization problems in geophysics, has already been demonstrated by previous authors. The work by Rothman (1985, 1986) on the residual statics estimation problem was the first published application to geophysical inverse problems. Later contributions have been made by Jakobsen, Mosegaard and Pedersen (1988) and Landa, Beydoun and Tarantola (1989). The experience of these authors and others who have tried to apply simulated annealing to geophysical inverse problems of a realistic size, is that the method is very difficult to use. The main problems are: to discover the best annealing temperature schedule or the 'critical' temperatures, and how many times the annealing should be performed in order to arrive at a useful solution. Consequently, a high degree of experimentation has been an important characteristic of simulated annealing work. The aim of our work has been to investigate the efficiency and accuracy of a recently developed implementation of simulated annealing applied to a realistic, seismic trace inversion problem. This method, which is called 'simulated annealing at constant thermodynamic speed' (Nulton and Salamon 1988; Andresen *et al.* 1988), replaces the random experimental approach with a systematic approach that takes advantage of statistical information about the model-algorithm system, acquired during the annealing. As a standard of reference in our investigation, we have used a more primitive, stochastic model optimization method, the so-called iterative improvement, when investigating the efficiency and accuracy of the implementation of simulated annealing.

Theory and Implementation

Simulated annealing is a statistical technique for finding near-optimal solutions to complex optimization problems. In this technique, the state ω of the system being optimized is identified with the state of a statistical mechanical system, the objective function $E(\omega)$ being minimized is identified with the physical energy, and the optimization process is controlled by a parameter T which can be identified with the physical temperature. The system to be optimized is allowed to 'equilibrate' by

applying a set of moves, i.e. a set of system perturbations, and accepting or rejecting the moves according to the Metropolis algorithm (Metropolis *et al.* 1953):

if $E_{\text{attempted}} \leq E_{\text{current}}$, accept the move;

if $E_{\text{attempted}} > E_{\text{current}}$, accept the move with probability

$$P_{\text{accept}} = \exp\left(-\frac{(E_{\text{attempted}} - E_{\text{current}})}{T}\right). \tag{1}$$

Under the following rather mild assumptions, it can be shown that the system tends towards 'thermal equilibrium' when iterated at any temperature (Hammersley and Handscomb 1964). (1) Any state of the system to be optimized can be reached from any other state of the system, using the prescribed move class. (2) There must be non-zero probability of staying in the current state in a given 'move'.

In thermal equilibrium at temperature T, the states ω of a hypothetical, large statistical ensemble of systems, identical to the considered system, are distributed according to the Boltzmann distribution

$$P_{\text{B}}(\omega) = \frac{\exp\left(-\frac{E(\omega)}{T}\right)}{Z(T)}, \tag{2}$$

where the partition function

$$Z(T) = \sum_{\omega} \exp\left(-\frac{E(\omega)}{T}\right). \tag{3}$$

Assuming that only one global minimum exists for E, the equilibrium ratio between the probability p_0 that the system is in the configuration ω_0, representing the global minimum for the objective function E_0, and the probability p_ω that the system is in any other state ω, corresponding to a value E_ω for the objective function, is

$$\frac{p_0}{p_\omega} = \exp\left(-\frac{E_0 - E_\omega}{T}\right), \tag{4}$$

showing that

$$\frac{p_0}{p_\omega} \geq 1, \tag{5}$$

and

$$\frac{p_0}{p_\omega} \to \infty \qquad \text{as} \qquad T \to 0. \tag{6}$$

The limit (6) can be explained as follows: consider the family of Boltzmann distributions (2), parametrized by the temperature parameter T, and consider the probabilities p_0 and p_ω that the equilibrated system is in the global minimum for the objective function (the ground state) and in the fixed state $\omega \neq \omega_0$, respectively. If

we consider the probabilities as functions of the temperature T, the limit (6) is valid. Since

$$\forall \omega: p_\omega \leq 1, \tag{7}$$

and

$$\sum_\omega p_\omega = 1, \tag{8}$$

for a system with a finite number of states (6) yields

$$p_0 \to 1 \qquad \text{as} \qquad T \to 0. \tag{9}$$

Therefore, optimization of our system can be achieved by attaining equilibrium at a low value of T.

The limit (9) suggests the following simple algorithm, which is known as 'simulated annealing' (Kirkpatrick, Gelatt and Vecchi 1983). (1) Distribute a number of copies of the model-algorithm system uniformly over the state space. This uniform distribution corresponds to a Boltzmann distribution at $T = \infty$. (2) Decrease T gradually from infinity to zero over a large number of steps, and let the system equilibrate approximately at each step. After this process is terminated, we have for each system, $p_0 \approx 1$.

A serious problem with this algorithm is that at low values of T, approximate equilibrium can only be attained in a large number of steps. It is therefore necessary to run the algorithm when the system is out of equilibrium, and in this case, (9) does not apply.

For non-equilibrium systems with ground state probability $p_0(T)$, we have in general,

$$p_0(T) \to \pi_0 \leq 1 \qquad \text{as} \qquad T \to 0, \tag{10}$$

where the probability π_0 depends on the way the temperature decreases with time. We must now find the optimal annealing schedule $T(t)$, satisfying the constraints

$$T(0) \gg 1 \tag{11}$$

and

$$T(t_{\max}) = 0, \tag{12}$$

where t is the time measured in number of Metropolis moves.

The optimal annealing schedule $T(t)$ should maximize π_0 for a given, finite number of iterations $t_{\max}$ (a given run time).

A solution has been proposed by Nulton and Salamon (1988). They suggest that the optimal annealing schedule keeps a constant difference between the (non-equilibrium) mean value $\langle E \rangle$ of the objective function of the system, and the mean value $\langle E \rangle_{\text{eq}}$ which the objective function would have had, if the system were in equilibrium at the considered temperature. The distance should be measured in units of the standard deviation $\sigma_E(T)$ of the fluctuating objective function, i.e.

$$\frac{\langle E\rangle - \langle E\rangle_{eq}}{\sigma_E(T)} = v, \tag{13}$$

where v is a constant named 'the thermodynamic speed'.

Combining (13) with the differential equation for simple, thermodynamic relaxation

$$\frac{d\langle E\rangle}{dt} = -\frac{\langle E\rangle - \langle E\rangle_{eq}}{\varepsilon(T)}, \tag{14}$$

where $\varepsilon(T)$ is the relaxation time of the system at temperature T, and the relationship between $\sigma_E(T)$ and the heat capacity $C(T)$ of the system at temperature T is given by

$$C(T) = \frac{\sigma_E^2(T)}{T^2}, \tag{15}$$

the following first-order differential equation in $T(t)$ is obtained

$$\frac{dT}{dt} = -\frac{vT}{\varepsilon(T)\sqrt{C(T)}}. \tag{16}$$

Simulated annealing using a schedule satisfying (16) is denoted 'simulated annealing at constant thermodynamic speed'. This implementation of simulated annealing is superior to previously used implementations (Salamon *et al.* 1988; Jakobsen *et al.* 1988).

The constant thermodynamic speed v in (16) is adjusted such that the temperature schedule resulting from the integration of (16) satisfies (11) and (12).

Andresen *et al.* (1988) describe methods by which $C(T)$ and $\varepsilon(T)$ can be estimated from statistical information about the system, collected during the annealing process. They suggest that a transition frequency matrix $\mathbf{Q}$, for attempted transitions, is formed during the annealing process. For each attempted transition, the values of the objective function, $E_{current}$ and $E_{attempted}$, for the current model and the trial model respectively, are saved. The ijth element of the current $\mathbf{Q}$ can then be determined since

$$Q_{ij} = \frac{n_{ij}}{\sum_j n_{ij}}, \tag{17}$$

where n_{ij} is the total number of attempted transitions (since the first iteration) between models having values of E between E_i and $E_i + \delta E$, and models for which E lies between E_j and $E_j + \delta E$. Here, δE is the constant difference between two successive, preselected levels, E_k and E_{k+1} of the objective function. A temperature dependent matrix $\mathbf{G}(T)$ can now be formed by multiplying elements of $\mathbf{Q}$ corresponding to $j > i$ with Boltzmann factors

$$\exp\left(-\frac{(E_j - E_i)}{T}\right), \tag{18}$$

and adjusting the diagonal elements of $\mathbf{G}(T)$ so as to keep the row sums equal to 1.

$\mathbf{G}(T)$ is a stochastic matrix (its row sum is equal to 1), and its largest eigenvalue is therefore 1. Let $\mathbf{p}$ be the corresponding, normalized eigenvector.

The temperature-dependent variance $\sigma_E^2(T)$ of the fluctuating objective function can now be estimated as

$$\sigma_E^2(T) = \langle E^2(T) \rangle - \langle E(T) \rangle^2, \tag{19}$$

where

$$\langle E^n \rangle = \sum_i (E_i)^n p_i. \tag{20}$$

The heat capacity is now calculated from (15).

If the second largest eigenvalue of $\mathbf{G}(T)$ is $\lambda_2(T)$, the relaxation time can be estimated as

$$\varepsilon(T) = -\frac{1}{\ln(\lambda_2)}. \tag{21}$$

Computer Simulations

We used simulated annealing at constant thermodynamic speed. We ran 50 system copies in parallel, each one having a unique starting model and a unique random number sequence.

Running several copies of the model-algorithm system enabled an assessment of the distribution of the computed model parameters. We used the dispersion of the final acoustic impedance values of the ten best models found, as a measure of the quality of the model optimization methods. This dispersion reflects the degree of non-uniqueness of the inverse problem, the influence of noise on the model estimate, and possible errors due to lack of convergence of the search algorithm.

In the present implementation of seismic trace inversion, the subsurface model was parametrized by the two-way traveltimes of the reflectors, and the acoustic impedances of the homogeneous layers between the reflectors. In the simulated annealing optimization, each move in the model space consisted of perturbing a randomly selected two-way traveltime or reflection coefficient, which obeyed the constraints imposed by the *a priori* information. Synthetic seismic traces were modelled by convolving subsurface reflectivities with a known, highly oscillatory wavelet, and the objective function used in the optimization was the energy of the difference between the true data and the modelled data. After the model optimization process, the surface acoustic impedance and the reflection coefficient series were mapped into acoustic impedance as a function of two-way traveltime.

Test Based on Real Well Data

The simulated annealing inversion method was compared with iterative improvement by solving a trace inversion problem in which the data consisted of a synthetic trace generated from well data. In the inversion we searched for an acoustic impedance function that was stepwise constant in depth and had a limited number of discontinuities, i.e. a 'blocked' impedance function.

Iterative improvement is one of the well-established traditional optimization

methods, and is therefore well suited as a standard of reference for the results obtained by simulated annealing. The iterative improvement method is very simple: a random move (obeying the move class specifications) in the model space is accepted only if it results in an improvement (a decrease) in the value of the objective function. It should be noted that iterative improvement is a local optimization method, and hence it is likely to converge towards local minima when used on the considered, global optimization problem.

The subsurface model used in the generation of the test data was derived from the sonic and the density logs of the onshore well Løgumkloster-1, situated in South Jutland in Denmark. Based on calibrated velocity and density logs, and a seismic wavelet extracted from a seismic line close to the well, a synthetic trace was generated (Fig. 2), using the detailed impedance function from the well (Fig. 1).

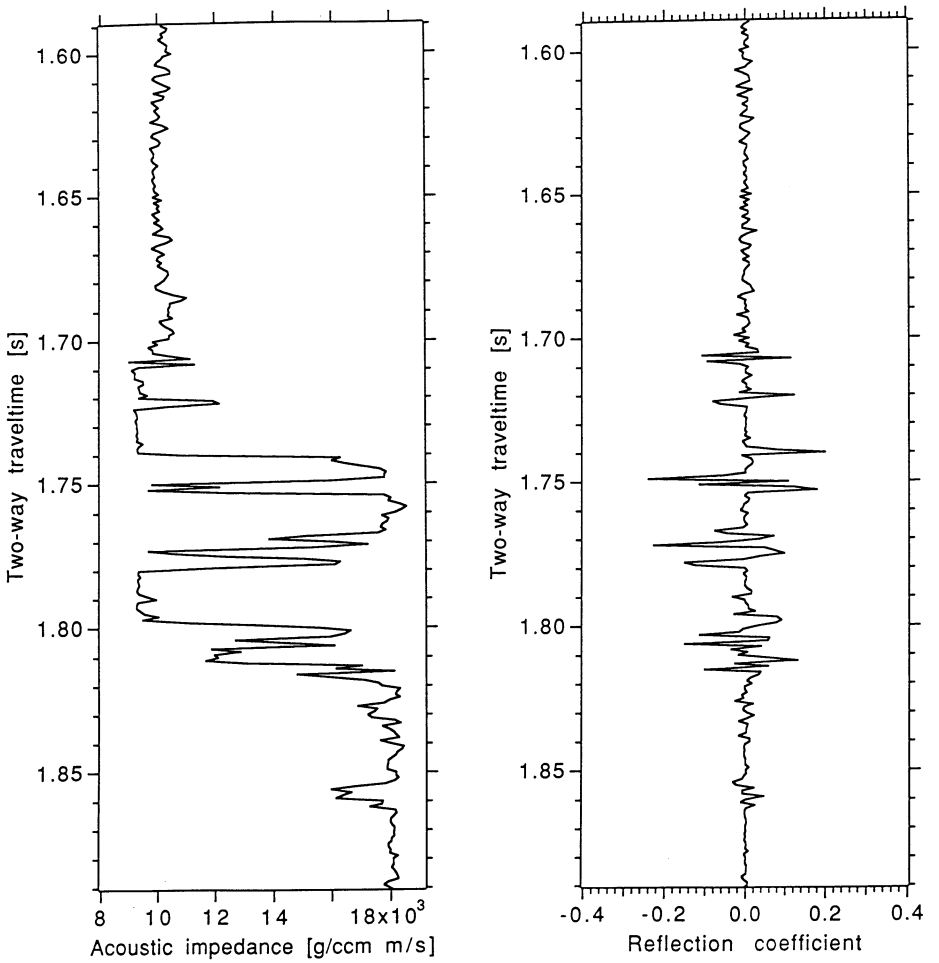

FIG. 1. The computed acoustic impedance (left) and reflection coefficients (right) versus two-way traveltime in the target zone.

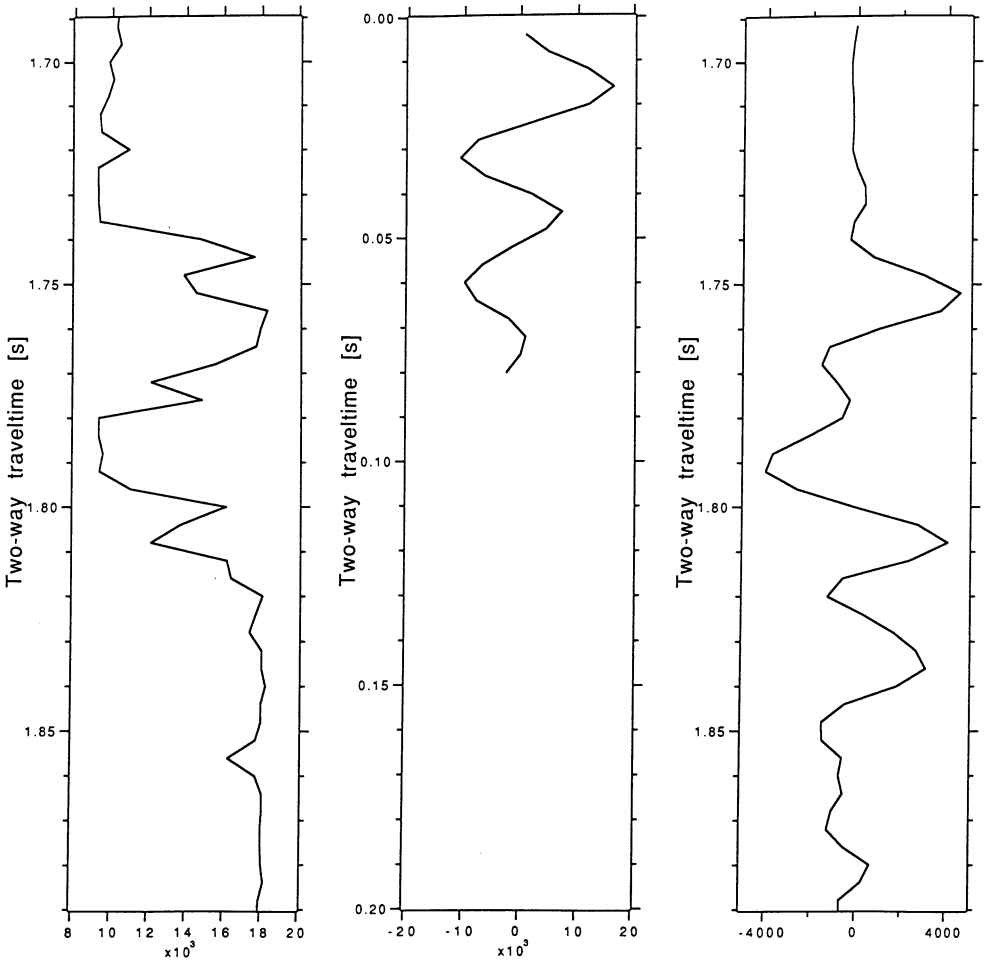

FIG. 2. The detailed acoustic impedance for the target zone (left), the estimated wavelet (centre), and the synthetic trace generated from the impedance log (right).

A manually blocked model for the Zechstein sequence is shown in Fig. 3. It is seen that ten homogeneous layers are sufficient to approximate the actual impedance log in this case. However, in order to simulate a realistic situation, we over-parametrized our model, assuming that it consisted of 15 homogeneous layers. The goal of the inversion was to reproduce approximately the well data from the seismic data and from the prior geological knowledge, without using the well log information.

The *a priori* information in the considered inverse problem is assumed to be sparse. The reflection coefficients are only known to be between -0.4 and 0.4, and there are no restrictions on the two-way traveltimes for the layer interfaces, except that they fall within the considered target zone.

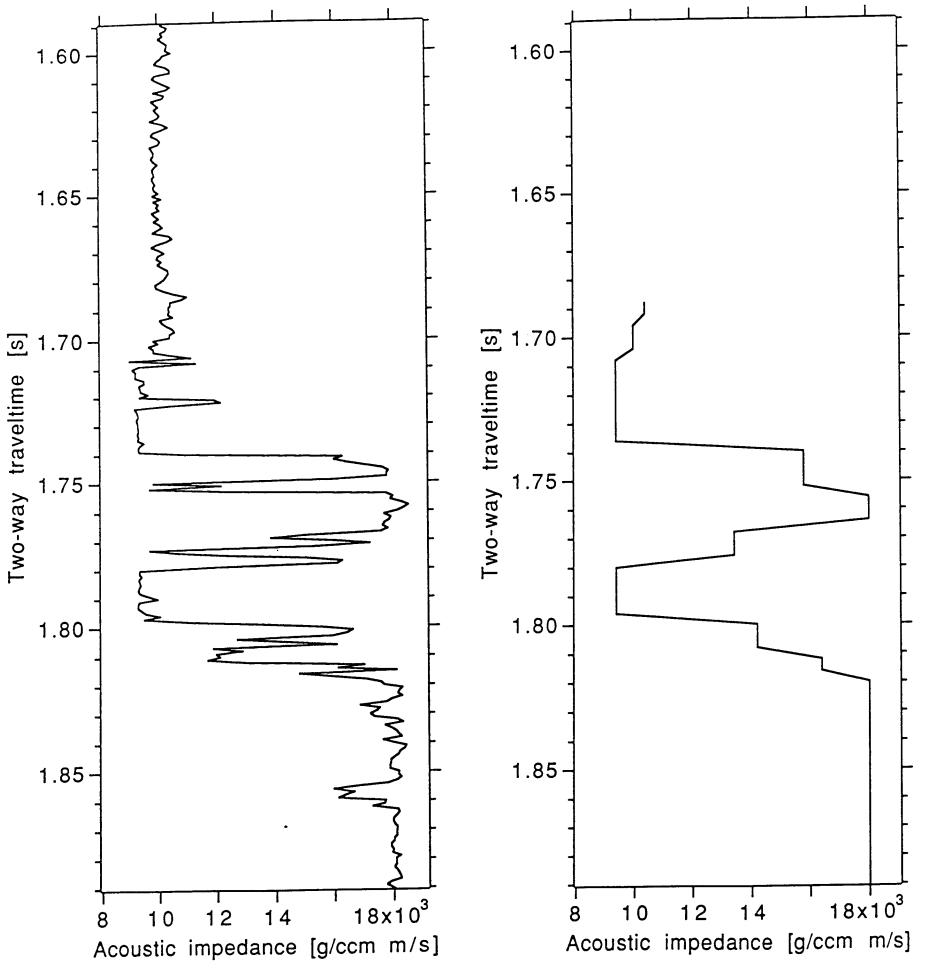

FIG. 3. The computed acoustic impedance (left) and the interpreter's blocking of the acoustic impedance (right) for the well Løgumkloster-1, resulting in a ten-layer model in the target zone.

The sparse prior information leads to multiple local minima for the objective function, typically appearing at parameter values representing cycle skips.

A 200 ms target zone from the synthetic reflection data set is inverted for a 15-layer impedance model. An important exploration problem to be solved in this area is to discriminate between salt layers and porous carbonate layers, on the basis of their acoustic impedance. In the data set used, a salt layer is found in the interval between 1.780 and 1.790 s, whilst the interval between 1.800 and 1.815 s is occupied by a porous carbonate.

An equal number of iterations were allocated to the two methods. First, 50 iterative improvement runs were performed, starting at different points, randomly

distributed in the model space. The individual runs were terminated when no significant changes in the value of the objective function had taken place within 500 iterations. The purpose of this terminating criterion was to optimize the use of the iterative improvement technique so that no time should be wasted by iterating after convergence to a local minimum had occurred. A total of 47 800 iterations were performed in this way.

Secondly, 50 simulated annealing runs were performed, all using the same number of iterations (namely 956), so that the total number of iterations was 47 800. Hence, no attempt was made to optimize the number of runs (or, equivalently, the number of iterations per run) performed by the annealing algorithm with the 47 800 iterations.

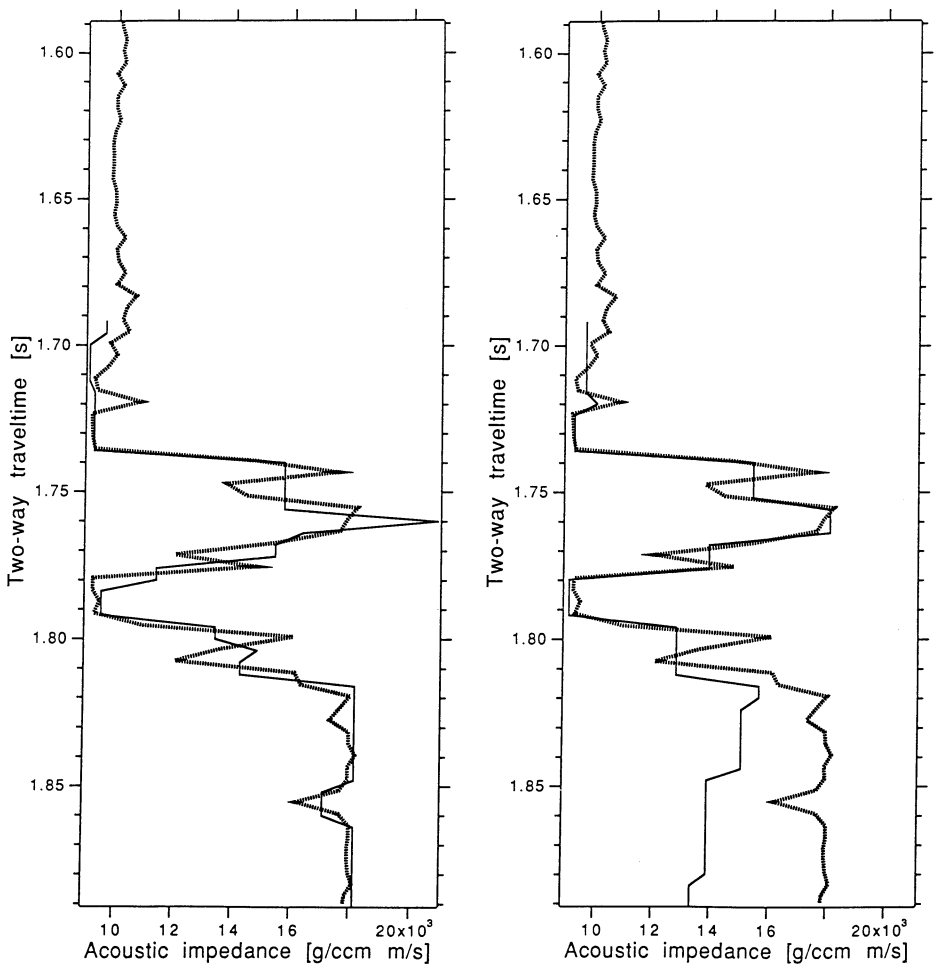

FIG. 4. The best acoustic impedances found by simulated annealing (left), and by iterative improvement (right). The true model is shown dotted.

In order to assess the performance of the two considered model optimization methods, we have chosen to display the ten best models obtained from each method, together with the true well data.

In this example, the chosen, blocked model parametrization is incapable of explaining all the data, due to the convolutional noise generated by the fine detail in the logs. This means that even the optimal solution has a finite, positive error energy. The best result obtained by simulated annealing is, however, in very good agreement with the well data (Fig. 4). In contrast to this, the best model obtained by iterative improvement fails to resemble the well data in the deeper part of the target zone. In particular, salt and porous carbonate are unlikely to be distinguished by inspection of the iterative improvement results. The near-optimal simulated annealing

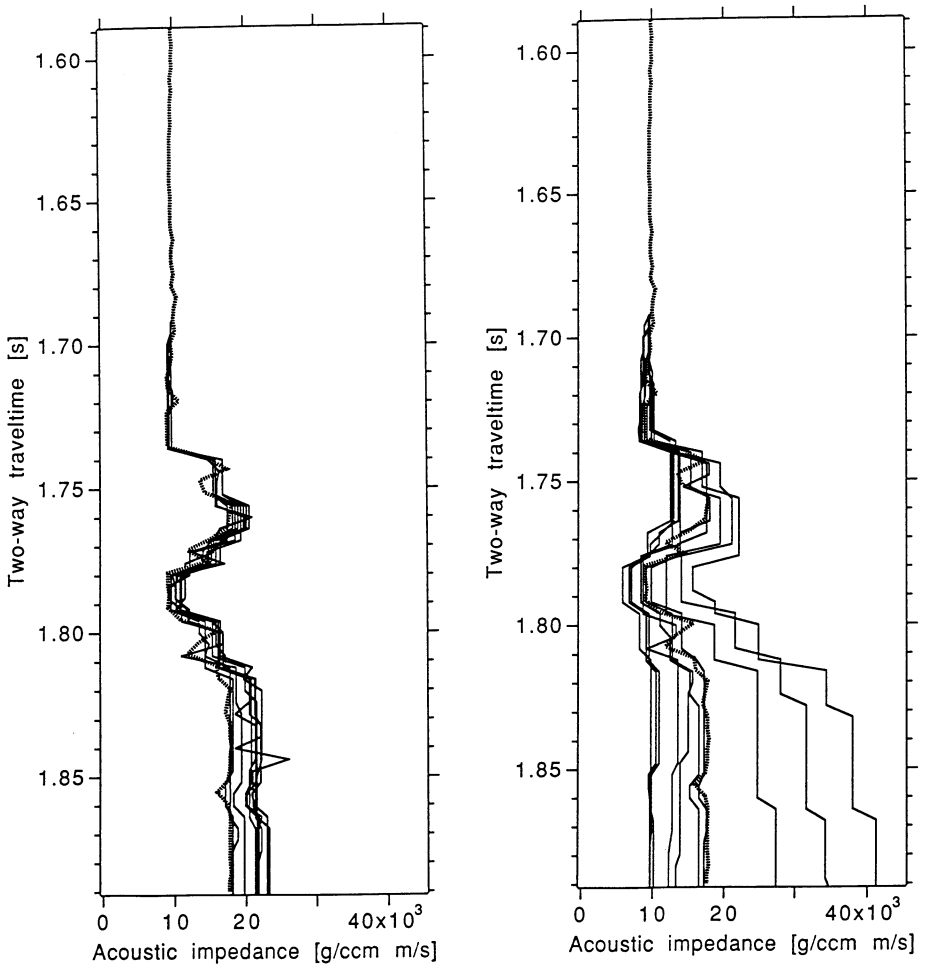

FIG. 5. The ten best impedance models obtained by simulated annealing (left) and by iterative improvement (right). The true model is shown dotted.

solution assigns separate impedance values to the two rock types, and hence enables an unambiguous rock identification.

To illustrate the uncertainties in the inversion procedures, the ten best computed models from the two methods are shown in Fig. 5. For the simulated annealing method, all the models shown are close to the true model, and the estimated impedances in the salt layer and in the porous carbonate layer are clearly separated. The best models from the iterative improvement method exhibit a gross divergence with increasing two-way traveltime, indicating that little confidence can be attached to the result of iterative improvement in this example.

Conclusion

We have investigated the use of a global, stochastic inverse method, i.e. simulated annealing, to solve the seismic trace inversion problem. The inverse problem was formulated as a search for a weakly constrained, blocked impedance function. The seismic data used were generated synthetically by convolving a known, highly oscillatory wavelet with the true reflectivity function from an onshore well located in South Jutland, Denmark. Hence, convolutional noise from a large number of thin layers in the subsurface was present in the data.

The simulated annealing algorithm employed in the present study was based on recent improvements by Nulton and Salamon (1988) and Andresen *et al.* (1988), in which statistical information about the system to be optimized is used to improve the performance of the algorithm. A traditional, stochastic model optimization method, iterative improvement, was used as a standard of reference when investigating the accuracy of the simulated annealing approach. 50 copies of the system, differing only in the starting models and the random sequences used, were run in parallel for both types of optimization. The dispersion of the ten best models obtained, and the similarity between these models and the true model were used as a measure of the performance of the algorithms.

The result of the investigation was that the best solutions found by the improved simulated annealing schedule displayed a significantly lower dispersion than the solutions found by iterative improvement, using the same total number of iterations. Furthermore, the best simulated annealing solutions were closest to the true model. The significance of these results was emphasized by the fact that only for iterative improvement was an attempt made to optimize the computations with respect to the number of iterations performed in each run.

The particular exploration problem of discriminating between salt and low porosity carbonate on the basis of the acoustic impedances of these rocks was solved by the simulated annealing approach, but remained unsolved by the iterative improvement algorithm.

Acknowledgements

This work was sponsored by The Danish Ministry of Energy under contract No. 1313/87-12. Klaus Mosegaard was partly sponsored by The Danish Natural Science Research Council. We thank Dr Daniel Rothman for his criticism and valuable

suggestions. We are indebted to Dr Jacob M. Pedersen, Ødegaard & Danneskiold-Samsøe, for assistance in the processing of our data.

REFERENCES

ANDRESEN, B., HOFFMANN, K.H., MOSEGAARD, K., NULTON, J.D., PEDERSEN, J.M. and SALAMON, P. 1988. On lumped models for thermodynamic properties of simulated annealing problems. *Journal de Physique* **49**, 1485–1492.

COOKE, D.A. and SCHNEIDER, W.A. 1983. Generalized linear inversion of reflection seismic data. *Geophysics* **48**, 665–676.

HAMMERSLEY, J.M. and HANDSCOMB, D.C. 1964. Monte Carlo Methods. In: Monographs on Statistics and Applied Probability. D.R. Cox and D.V. Hinkley (eds). Chapman and Hall.

JAKOBSEN, M.O., MOSEGAARD, K. and PEDERSEN, J.M. 1988. Global model optimization in reflection seismology by simulated annealing. In: *Model Optimization in Exploration Geophysics* **2**, p. 361. Proceedings of the 5th International Mathematical Geophysics Seminar, Berlin 1987, A. Vogel (ed.). Friedr. Vieweg & Son, Braunschweig, Wiesbaden.

KIRKPATRICK, S., GELATT, C.D. and VECCHI, M.P. 1983. Optimization by simulated annealing. *Science* **220**, 671–680.

LANDA, E., BEYDOUN, W. and TARANTOLA, A. 1989. Reference velocity model estimation from prestack waveforms: coherency optimization by simulated annealing. *Geophysics* **54**, 984–990.

METROPOLIS, N., ROSENBLUTH, A.W., ROSENBLUTH, M.N., TELLER, A.H. and TELLER, E. 1953. Equation of state calculations by fast computing machines. *Journal of Chemical Physics* **1**, 1087–1092.

NULTON, J.D. and SALAMON, P. 1988. Statistical mechanics of combinatorial optimization. *Physical Review* A **37**, 1351–1356.

ROTHMAN, D.H. 1985. Nonlinear inversion, statistical mechanics, and residual statics estimation. *Geophysics* **50**, 2797–2807.

ROTHMAN, D.H. 1986. Automatic estimation of large residual static corrections. *Geophysics* **51**, 332–346.

SALAMON, P., NULTON, J.D., HARLAND, J.R., PEDERSEN, J.M., RUPPEINER, G. and LIAO, L. 1988. Simulated annealing with constant thermodynamic speed. *Computer Physics Communications* **49**, 423–428.