Seismic traveltime tomography: A simulated annealing approach

Article *in* Physics of The Earth and Planetary Interiors · April 2000

DOI: 10.1016/S0031-9201(99)00157-0

1 author:

Zoltán Wéber
Institute of Earth Physics and Space Science (EPSS)
60 PUBLICATIONS **1,202** CITATIONS

ELSEVIER

Physics of the Earth and Planetary Interiors 119 (2000) 149–159

PHYSICS OF THE EARTH AND PLANETARY INTERIORS

www.elsevier.com/locate/pepi

Seismic traveltime tomography: a simulated annealing approach

Zoltán Wéber *

Seismological Observatory of the Hungarian Academy of Sciences, Meredek u. 18, H-1112 Budapest, Hungary

Received 13 November 1997; received in revised form 1 November 1998; accepted 14 November 1998

Abstract

Seismic traveltime tomography involves finding a velocity model that minimizes the error energy between the measured and the theoretical traveltimes. When solving this nonlinear inverse problem, a local optimization technique can easily produce a solution for which the gradient of the error energy function vanishes, but the energy function itself does not take its global minimum. Other methods such as simulated annealing can be applied to such global optimization problems. The simulated annealing approach to seismic traveltime tomography described in this paper has been tested on synthetic as well as real seismic data. It is shown that unlike local methods, the convergence of the simulated annealing algorithm is independent of the initial model: even in cases of virtually no prior information, it is capable of producing reliable results. The method can provide a number of acceptable solutions. When prior information is sparse, the solution of the global optimization can be used as an input to a local optimization procedure, such as, e.g., simultaneous iterative reconstruction technique (SIRT), producing an even more accurate result.

Keywords: Seismic tomography; Global optimization; Simulated annealing

1. Introduction

Automatic tomographic inversion schemes for the reconstruction of subsurface structures from seismic traveltime data are used more and more frequently. An analysis of the principles behind the most popular methods reveals that these techniques all belong to a category called local optimization methods. The most popular local optimization methods are the algebraic reconstruction techniques (ART), simultaneous iterative reconstruction techniques (SIRT), damped least squares, the conjugate gradients methods, and others (Herman et al., 1973; Gordon, 1974; Dines and Lytle, 1979; Peterson et al., 1985; van der Sluis and van der Vorst, 1987; Phillips and Fehler, 1991; and many others). The principal problem of these methods is well known. They require a good starting model since they look for a solution in the close neighborhood of the starting model. Since the error energy function to be minimized in the course of a tomographic inversion may have many points with vanishing gradient, if the trial solution is too far from the global minimum, the method may stop the inversion process at any point where the error energy function does not change significantly.

One of the alternatives to these methods is the systematic exploration of the model space, in which each point in the model space is searched sequen-

* Fax: +36-1-3193385; e-mail: weber@seismology.hu

0031-9201/00/$ - see front matter
PII: S0031-9201(99)00157-0

tially. This, however, is an impossible task because our model space is extremely large.

The solution to the above-described problems is to employ a global optimization method. Global optimization problems are much more difficult to solve than local optimization problems, since the local geometry of the cost function surface in the model space does not directly contain information about the direction to follow in a search for the global minimum. Therefore, a stochastic optimization technique has to be applied in a global search for the optimal subsurface model. The simulated annealing method belongs to this category.

In order to illustrate that the application of global optimization techniques in traveltime tomography is desirable, a simple synthetic inversion problem was investigated in detail (Fig. 1). The true velocity model with the source–receiver configuration is depicted in Fig. 1a. The velocity values of v_1 and v_2 are to be estimated by minimization of the root mean square (rms) error between the ''measured'' and the theoretical traveltimes. Since several head waves are generated during this simulated experiment, the problem can be solved. Fig. 1b shows the misfit function, i.e., the rms data error, as a function of v_1 and v_2. The arrows in the figure represent the directions and magnitudes of the negative gradient vectors of the error surface. Since the misfit function does not depend on v_2 when $v_1 \geq v_2$, it has a long valley with vanishing gradient at about $v_1 = 1080$ m/s. It is obvious that a local optimization method easily falls into that valley, if the starting model is somewhere around the point *A* in Fig. 1b. In this case, v_1 becomes equal to the average velocity and v_2 can be equal to any value not greater than v_1. On the other hand, a global optimization method, such as simulated annealing, has every chance to get out of the valley and to find the global minimum. I think that situations similar to the one outlined above can frequently arise in more complicated problems with thousands of velocity cells and noisy observations.

Since the work of Kirkpatrick et al. (1983), the simulated annealing method has been used in many parameter optimization problems (van Laarhoven and Aarts, 1987). 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 large residual statics estimation problem was the first published geophysical application of the simulated annealing procedure. A similar problem was solved by Vasudevan et al. (1991). Mosegaard and Vester-

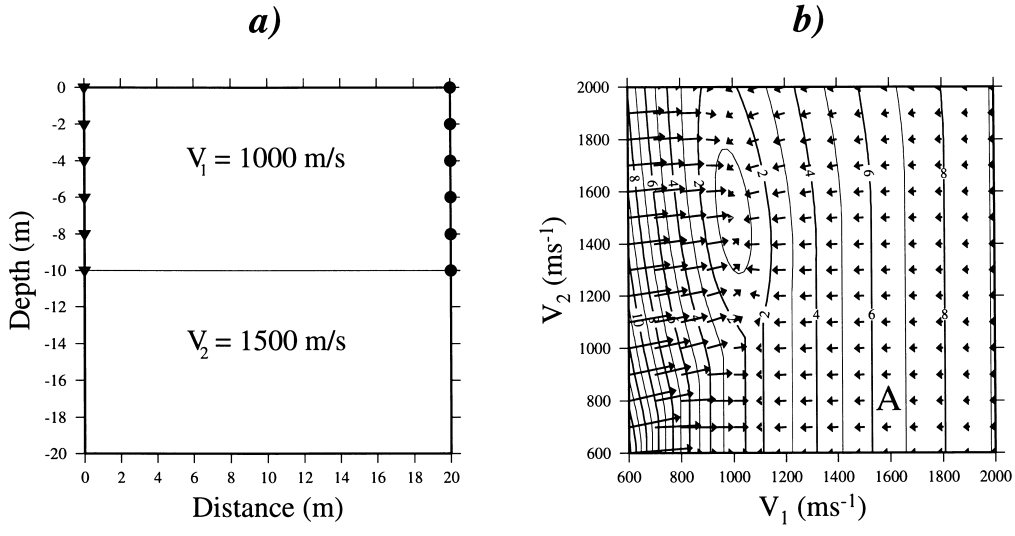

Fig. 1. A simple synthetic inversion problem. (a) The true velocity model. Source and receiver positions are indicated by triangles and circles, respectively. (b) Contour plot of the rms data error as a function of v_1 and v_2. Contour values are in milliseconds. The arrows represent the negative gradient vectors of the error surface. A local optimization method easily falls into the valley at about $v_1 = 1080$ m/s, if the starting model is somewhere around the point *A*.

gaard (1991) and Sen and Stoffa (1991) successfully applied this technique to the one-dimensional seismic waveform inversion problem. Landa et al. (1989) and Vestergaard and Mosegaard (1991) pointed out that the annealing method can produce reliable results efficiently in case of two-dimensional seismic waveform inversion problems, even when prior information is sparse. Pullammanappallil and Louie (1994) used the method for inversion of first arrival times of reflection experiments.

In this paper we apply the simulated annealing method to seismic transmission traveltime tomographic inversion and compare its solutions to those produced by SIRT, a well known local optimization procedure.

2. The algorithms

2.1. Raypath calculation

Ray tracing is the most time-consuming step of nonlinear seismic tomography. Since usually thousands of rays must be traced for a typical tomographic inversion, saving computer time and memory plays an important role in selecting the appropriate forward modeling algorithm.

Shortest path ray tracing (SPR) has many potential benefits over other forward modeling techniques. It works well with block or otherwise discrete velocity models, it is fairly fast when used to trace many rays, and it always finds the global minimum traveltime raypath through a network that represents the earth (Moser, 1991). Diffracted raypaths, headwaves and paths to shadow zones can be found correctly. There are also no restrictions to the complexity or the dimensionality of the velocity model.

However, when accurate results are needed, the method requires vast memory and intensive calculation. In order to achieve the appropriate accuracy with as little computer time as possible, we have chosen an improved version of SPR proposed by Wéber (1995) as a tool for raypath computation.

2.2. Simulated annealing

Simulated annealing method models a physical process in which a solid in a heat bath is warmed by increasing the temperature. This is followed by slow cooling until the global minimum energy state is reached, where it forms a crystal. Metropolis et al. (1953) introduced a simple algorithm that can be used to simulate a collection of atoms in equilibrium at a given temperature. In each step of this algorithm an atom is given a small random displacement and the resulting change in the energy of the system, ΔE, is computed. If $\Delta E \leq 0$, the displacement is accepted and the new configuration is used as the starting point of the next step. If $\Delta E > 0$, the new configuration is accepted with a probability of $P(\Delta E) = \exp(-\Delta E/kT)$. By repeating the basic step many times, thermal equilibrium at temperature T is simulated. The choice of $P(\Delta E)$ has the consequence that the system evolves into Boltzmann distribution.

Using the cost function (which, in our problem, is the rms error between the measured and the theoretical traveltimes) in place of the energy and defining configurations by a set of parameters (velocity values in our case), it is straightforward with the Metropolis procedure to generate configurations for a given optimization problem (e.g., tomographic problem) at some effective temperature (Kirkpatrick et al., 1983; Bohachevsky et al., 1986). This temperature is simply a control parameter in the same units as the cost function. Simulated annealing consists of first ''melting'' the system being optimized at a high effective temperature, then decreasing the temperature slowly until the system ''freezes'' and no further changes occur. At each temperature, the simulation must proceed long enough for the system to reach a steady state. A simple transition from one state of the system to another one is called a simulation step. The starting temperature, the sequence of temperatures, the number of the simulation steps at each temperature, and the stopping criterion is called the annealing schedule.

2.3. Annealing schedule

The simulated annealing algorithm converges with probability 1 to the global minimum of the cost function if for each value of the temperature, T_k, an infinite number of simulation steps is generated and $\lim_{k \to \infty} T_k = 0$ (van Laarhoven and Aarts, 1987). In any implementation of the algorithm, asymptotic convergence can only be approximated.

We resort to an implementation in which a finite number of simulation steps, L_k, is generated at decreasing values of the temperature, T_k. The annealing schedule is then specified as follows.

2.3.1. Initial value of the temperature

The initial value of T, i.e., T_0, should be as small as possible but in such a way that virtually all simulation steps are accepted. To estimate the optimal value of T_0, we use a procedure based on the one developed by White (1984). It essentially consists of computing the values of the cost function for a predefined number of system configurations randomly chosen around the initial configuration. Then the initial value of the temperature is calculated as:

$$T_0 \geq \sqrt{\overline{E^2(\mathbf{v})} - \overline{E(\mathbf{v})}^2},$$

where $E(\mathbf{v})$ is the cost function and $\mathbf{v}$ represents the system configuration. In other words, T_0 is chosen to be greater than or equal to the standard deviation of the previously calculated cost values.

2.3.2. Decrement of the temperature

The decrement should be chosen such that a few number of simulation steps suffice to re-establish quasi-equilibrium at each temperature T_k. We use the simple $T_{k+1} = 0.99T_k$ decrement rule.

2.3.3. Number of simulation steps at T_k

In our implementation L_k is determined such that for each value of T_k a minimum number of simulation steps, $A_{\min}$ is accepted, where $A_{\min}$ is a predefined fix number. However, as T_k approaches 0, acceptance probability decreases and eventually the value of L_k reaches infinity. Consequently, L_k is not allowed to exceed some constant $L_{\max}$ to avoid extremely large number of simulation steps at low values of T_k.

2.3.4. Stopping criterion

Let U denote the number of the consecutive temperatures for which the number of the accepted simulation steps is zero. If U exceeds a predefined number, $U_{\max}$, the annealing procedure stops. In other words, the procedure stops if the system configuration is not altered for $U_{\max}$ number of consecutive temperatures, i.e., if the configuration remains unchanged for $U_{\max} \cdot L_{\max}$ number of consecutive simulation steps.

2.4. Implementation

According to the above discussion the implementation of the simulated annealing can be summarized as follows.

Assume $E(\mathbf{v})$ is the rms difference between the measured and the theoretical traveltimes and it is to be minimized with $\mathbf{v}$ restricted to Ω, a subset of R^n, where n is the number of the unknown velocities. Initially a starting velocity model, $\mathbf{v}_0$, is chosen arbitrarily on Ω, and the initial value of the temperature, T_0, is determined as described above. Furthermore, let A_k denote the number of the accepted simulation steps for temperature T_k. Then the algorithm essentially comprises the following steps:

(1) Compute traveltimes through the initial velocity model $\mathbf{v}_0$. Determine the rms error $E(\mathbf{v}_0)$. Set $k = 0$, and $U = 0$.

(2) Set $L_k = 0$, and $A_k = 0$.

(3) The velocity model is altered randomly; each velocity value may remain unchanged or may decrease or increase by a predefined velocity step, Δv, with equal probability. After some two-dimensional smoothing we get a new model, $\mathbf{v}_1$. If $\mathbf{v}_1 \in \Omega$, we calculate $\Delta E = E(\mathbf{v}_1) - E(\mathbf{v}_0)$, otherwise a new $\mathbf{v}_1$ is generated. This step is called a *simulation step*. Increase L_k by 1.

(4) If $\Delta E \leq 0$, the simulation step is accepted unconditionally, i.e., set $\mathbf{v}_0 = \mathbf{v}_1$ and increase A_k by 1.

(5) If $\Delta E > 0$, the simulation step is accepted with probability:

$$P = \exp(-\Delta E/T_k).$$

Generate a uniformly distributed number ρ between 0 and 1. If $\rho < P$, the simulation step is accepted, i.e., set $\mathbf{v}_0 = \mathbf{v}_1$ and increase A_k by 1. This step provides for the conditional acceptance of models with a larger rms error. This gives the inversion the ability to escape out of local minima in its search for the global minimum.

(6) If $A_k < A_{\min}$ and $L_k < L_{\max}$, go to step 3.

(7) If $A_k = 0$, increase U by 1, otherwise set $U = 0$.

(8) If $U \leq U_{\max}$, update the temperature: $T_{k+1} = 0.99T_k$, and increase k by 1. Go to step 2.

(9) If $U > U_{\max}$, stop. The best solution is $\mathbf{v}_0$.

Throughout this paper 100 randomly chosen velocity models are used for the estimation of T_0, while the other annealing parameters take the following values: $\Delta v = 20$ m/s, $A_{\min} = 20$, $L_{\max} = 40$, and $U_{\max} = 100$.

The two-dimensional smoothing applied in each simulation step has important roles in the algorithm. Firstly, smoothing decreases the variance of the velocity distribution. The purpose of minimizing variance is to find the smoothest possible velocity image, i.e., the one closest to the mean velocity which still satisfies the data. The smoothest velocity model is the least likely to contain artifacts (Herman et al., 1973). Thus, in case of noisy data the application of a smoothing operator is unavoidable (Dines and Lytle, 1979). Secondly, smoothing can constrain the velocity values in those cells that are not hit by any rays at all. Finally, smoothing forces the simulated annealing algorithm to try to find the solution among only the smooth velocity distributions of the model space, thus making the procedure more efficient.

3. Experiments

In this section we present the results of our tests of the annealing optimization scheme on synthetic models. In addition we also compare its performance with a popular local optimization method called SIRT. A test on real data is also demonstrated.

A synthetic cross-hole data set was generated. The model is 50 m long and 50 m deep and it has a background velocity of 1000 m/s and a step-like high-velocity layer with velocity of 1300 m/s. The layer boundaries are superimposed on all of the following figures. We placed 51 seismic sources at the left edge of the model and 101 receivers at the surface and at the right edge of the model. Both the sources and the receivers were located equidistantly at 1 m intervals. This configuration produces 5150 rays with a broad range of raypath coverage over the model. In the course of the inversion, the model is divided into 1×1 m cells, i.e., there are 2500 unknown velocity values and 5150 traveltime observations.

In order to make the annealing algorithm more efficient, the synthetic raypaths are updated only when the temperature is updated. Then, at a given temperature, at each simulation step the traveltimes can be calculated very rapidly using the stored raypaths and the velocity values in the model cells. This approach presumes that in the course of the inversion, at a constant temperature, the velocity values do not change significantly. In this way the SPR algorithm, the most time-consuming step of the inversion procedure, has to be executed much less times than if the raypaths were updated at each simulation step.

In order to increase the general stability of the solutions, after each simulation step the randomly altered velocity values are smoothed by a 3×3 median filter. The median filter is chosen since it does not destroy high-contrast boundaries as an averaging smoother would and it is not affected by isolated spikes that occasionally occur in tomographic images, when some velocity cells are not crossed by any rays at all.

The inversion results are characterized by two error measures: percent of rms error between the final solution and the original model (σ_v) and the rms data residual (σ_t) which indicates how well the traveltime data are predicted by the solution. This latter misfit function is to be minimized by the inversion procedure.

Before presenting the results of the synthetic tests, we briefly explain how the SIRT algorithm is carried out. The raypath matrix is updated in each iteration step according to the current velocity distribution using the same SPR algorithm that is also used in the course of the simulated annealing procedure. After each iteration step, the velocity values are smoothed by a 3×3 median filter, and the variance of the current velocity model and the rms data error are calculated. The iteration process is stopped at the first point, after which the change in velocity variance is less than 1% of the current variance and the change in data error is less than 5% of the current error (Herman et al., 1973).

3.1. Good prior information

To simulate good prior information, a homogeneous starting model with velocity of $v_0 = 1000$ m/s is defined. We allow the velocity to vary between 500 and 1500 m/s, i.e., $\Omega = \{\mathbf{v} \in R^n | 500 \text{ m/s} \le v_i \le 1500 \text{ m/s}, \ 1 \le i \le n\}$. Figs. 2 and 3

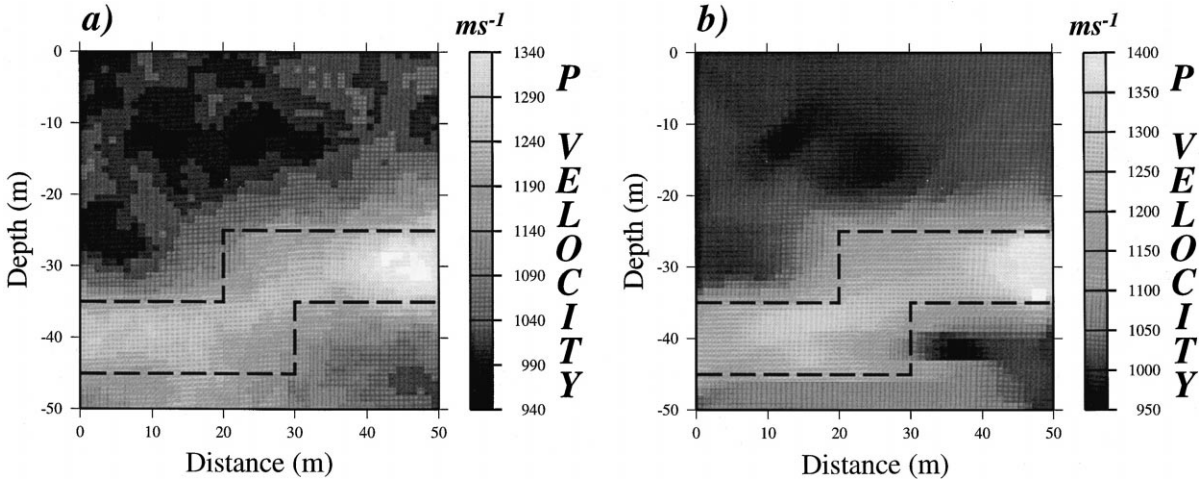

Fig. 2. Inversion results of synthetic experiments using noise-free data and good starting model. (a) Reconstructed image produced by the annealing scheme. (b) Reconstructed image calculated by SIRT.

show the results of the optimization process compared to those of SIRT for noise-free and noisy traveltime data, respectively. The error measures are summarized in Table 1. Noisy data means that random noise with standard deviation of 2 ms was added to the exact traveltimes resulting in about 5–10% average absolute relative data error.

For both the noise-free and the noisy traveltime data, the simulated annealing method reconstructs the high-velocity region and the background velocity well (Figs. 2a and 3a, respectively). The rms data error reduces from a value of about 5 to 0.49 ms in case of no data noise (Fig. 4a). From the figure we see that during the initial stages of the optimization, when the temperature is high, the rms data residual varies rapidly. This indicates that the annealing procedure tests numerous models and can easily escape from a local minimum. As the optimization pro-

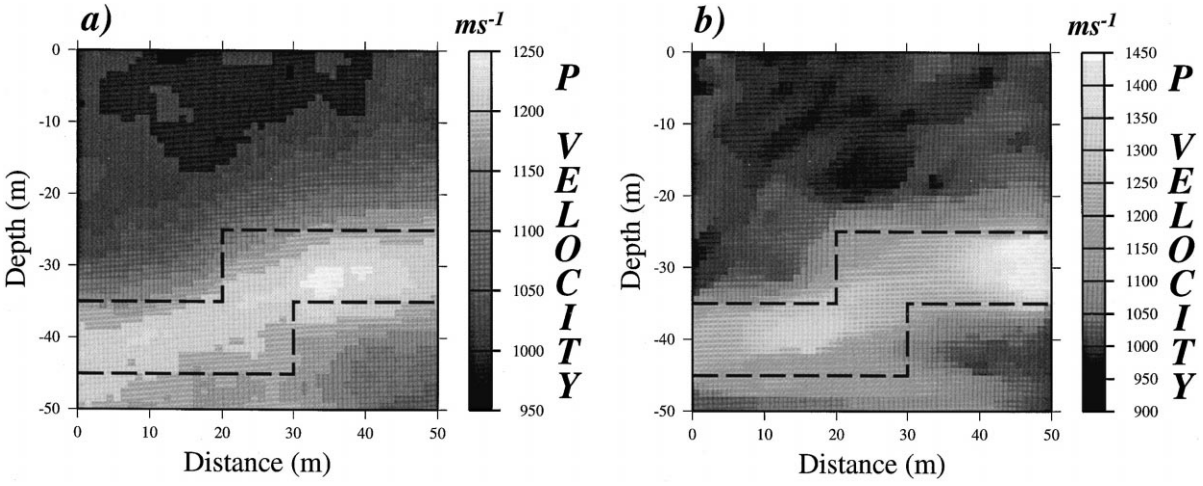

Fig. 3. Inversion results of synthetic experiments using noisy data and good starting model. Noisy data are generated by adding random noise with standard deviation of 2 ms to the exact traveltimes. (a) Reconstructed image produced by the annealing scheme. (b) Reconstructed image calculated by SIRT.

Table 1
Results of synthetic experiments

Inversion method →		SIRT		Annealing		Annealing + SIRT	
		σ_v (%)	σ_t (ms)	σ_v (%)	σ_t (ms)	σ_v (%)	σ_t (ms)
''good'' model	noise-free	6.35	0.20	8.37	0.49	7.05	0.20
	noisy	6.97	1.98	9.24	2.22	7.23	2.00
''bad'' model	noise-free	–	–	8.60	0.51	7.22	0.20
	noisy	–	–	9.38	2.20	7.43	2.00

gresses and the temperature becomes smaller, the probability of escaping out of local minima also decreases. This shows that the choice of T_0 and the cooling rate affect the convergence of the annealing process.

The inversion results calculated by SIRT are more accurate than those determined by the annealing algorithm: the contours of the high-velocity layer are much sharper (Figs. 2b and 3b) and the error measures have smaller values (Table 1). Even if the annealing results are used as starting models for SIRT, the resulting velocity sections are not superior to those produced by SIRT alone (Table 1). The detailed analysis of the solutions reveals that inside the high-velocity layer the values of σ_v are approximately equal for both the SIRT and the annealing procedure. Outside the layer, however, annealing yields less accurate result than SIRT does. After all, if we have good a priori information about the velocity distribution, it seems to be more economical to use a local optimization technique than a global one.

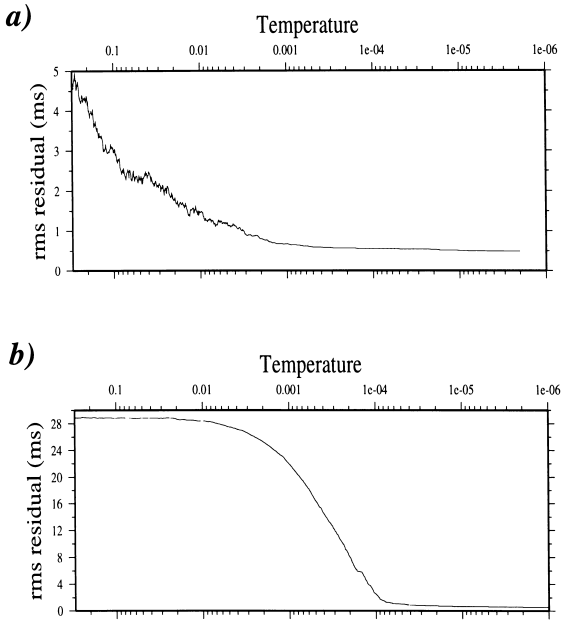

Fig. 4. The rms data residual variation during the optimization process for (a) noise-free data and good starting model and (b) noise-free data and poor initial model.

3.2. Sparse prior information

To simulate sparse prior information, a homogeneous starting model with velocity of $v_0 = 3000$ m/s is defined. We allow the velocity to vary between 500 and 5000 m/s, i.e., $\Omega = \{\mathbf{v} \in R^n | 500 \text{ m/s} \le v_i \le 5000 \text{ m/s}, \ 1 \le i \le n\}$. Figs. 5 and 6 show the results of the optimization process. The error measures are summarized in Table 1.

SIRT, the local optimization method, fails to converge and has not produced any valuable results. It is because the convergence of a local optimization method strongly depends on the initial model. Since the data misfit function to be minimized may have many points with vanishing gradient, if the starting model is too far from the global minimum, the method may stop the inversion process at any point where the misfit function does not change significantly.

On the other hand, global optimization methods, such as simulated annealing, find the global minimum independently of the initial model. Both for the noise-free and the noisy traveltime data, the method again reconstructs the high-velocity region and the background velocity well (Figs. 5a and 6a, respectively). The rms data error reduces from a value of about 29 to 0.51 ms in case of no data noise (Fig. 4b). The results are very similar to those obtained by the good starting model. The random perturbations

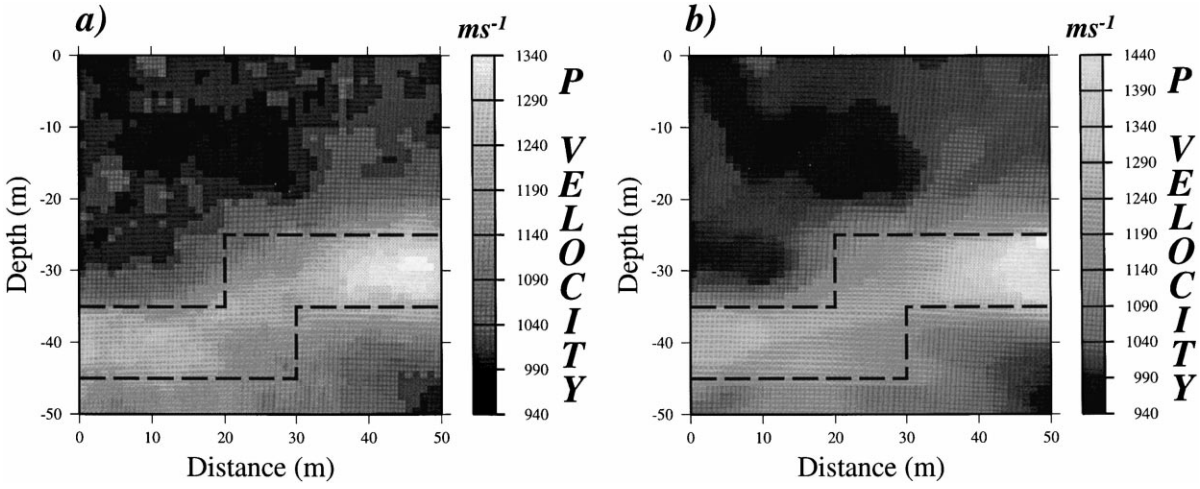

Fig. 5. Inversion results of synthetic experiments using noise-free data and poor starting model. (a) Reconstructed image produced by the annealing scheme. (b) Reconstructed image produced by SIRT using the velocity distribution in (a) as the starting model.

used during the annealing scheme make it independent of the initial model. Annealing destroys any order in the starting model within the first few iterations.

When the velocity model was altered randomly, a velocity step of 20 m/s was used, i.e., in each cell the velocity might have any of 226 distinct values. Since the number of cells is 2500, the model space consisted of 226^{2500} discrete models. On the other hand, in the course of the annealing procedure the temperature had about 1500 distinct values. At each temperature, the number of simulation steps was between 20 ($A_{\min}$) and 40 ($L_{\max}$). In other words, the annealing algorithm visited about 45,000 models and it was enough to find a fairly good solution in the huge model space.

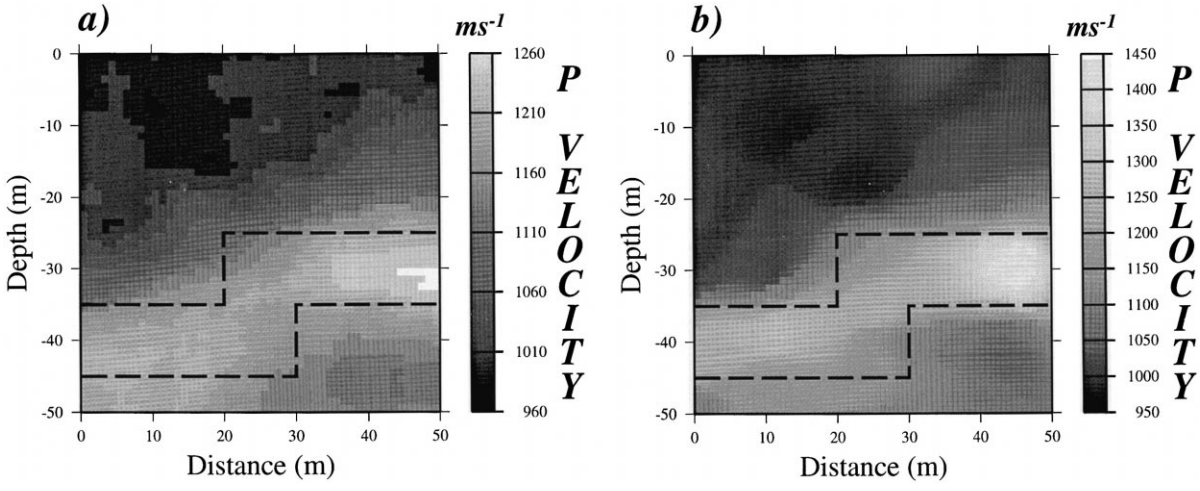

Fig. 6. Inversion results of synthetic experiments using noisy data and poor starting model. Noisy data are generated by adding random noise with standard deviation of 2 ms to the exact traveltimes. (a) Reconstructed image produced by the annealing scheme. (b) Reconstructed image produced by SIRT using the velocity distribution in (a) as the starting model.

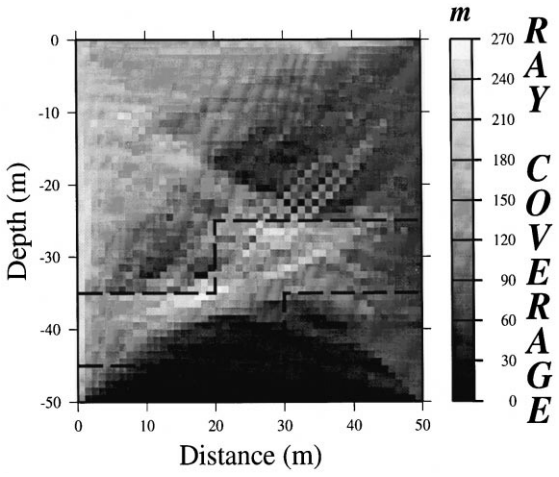

Fig. 7. Ray coverage for the velocity model depicted in Fig. 5a.

Being very close to the global minimum, the annealing results can be used as starting models to a local optimization method producing even more accurate results. The final models illustrated in Figs. 5a and 6b are produced by SIRT using the velocity distributions depicted in Figs. 5a and 6a as initial models, respectively. The contours of the high-velocity layer become much sharper allowing the interpreter to mark out the layer boundaries. It should be noted again that without a global optimization method we would have never got these accurate velocity distributions.

The simulated annealing method provides a number of acceptable solutions that have comparable rms data error near the global minimum. One may think that it allows us to estimate the uncertainty associated with the velocity values obtained using statistical analysis tools. Unfortunately, it is not true because there might be unconstrained regions in the model that are not perturbed and hence have a small deviation. However, the so-called ray coverage is a

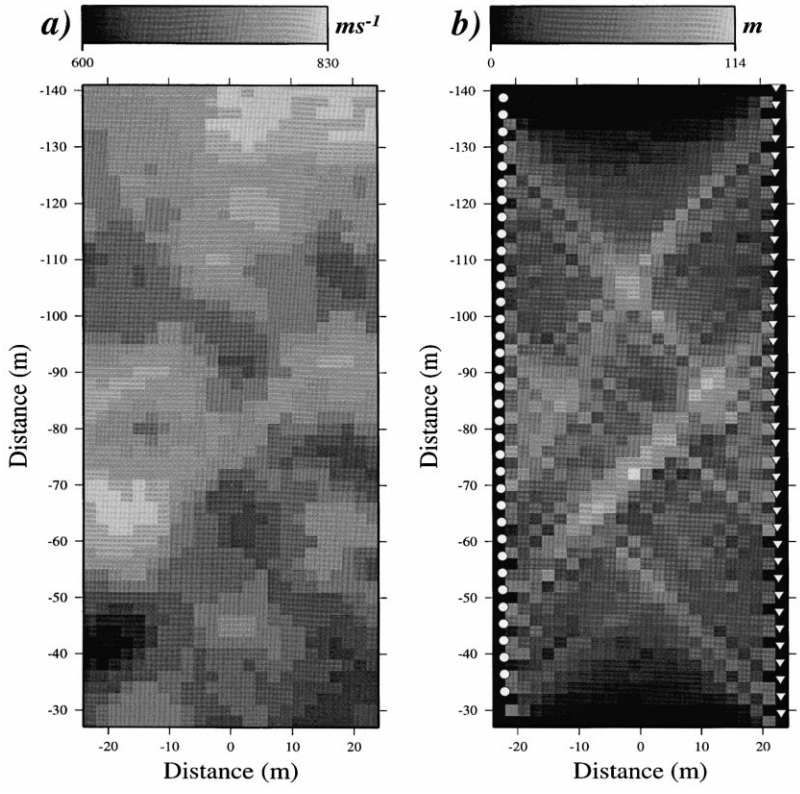

Fig. 8. (a) Reconstructed velocity image using a real data set measured in a coal mine. (b) Ray coverage associated with the velocity model in (a). Source and receiver positions are also indicated by triangles and circles, respectively.

valuable tool for estimating the uncertainty of the solutions qualitatively. The ray coverage of a given velocity cell is defined as the total length of ray segments intersecting that cell. Low value of the ray coverage in a cell usually involves greater uncertainty of the velocity value in that cell. For example, the ray coverage associated with the annealing result of Fig. 5a suggests that the velocity values at the bottom of the model have the greatest uncertainties (Fig. 7). Thus, the velocity values at and near the bottom of the high-velocity layer are uncertain. Indeed, the layer boundary in that region of the model is rather indefinite even after the application of SIRT (Fig. 5b).

3.3. Real example

The proposed method has also been tested on real data measured in a Hungarian coal mine. The resulting velocity distribution and ray coverage are depicted in Fig. 8. The seismic sources and receivers are placed on the right and left edge, respectively, of the investigated section, as shown in Fig. 8b. In the course of the inversion the model is divided into 2×2 m cells. A homogeneous starting model with velocity of $v_0 = 800$ m/s is defined. We allow the velocity to vary between 500 and 1100 m/s, i.e., $\Omega = \{\mathbf{v} \in R^n | 500 \text{ m/s} \leq v_i \leq 1100 \text{ m/s}, \ 1 \leq i \leq n\}$. Because the section is almost homogeneous geologically, the reconstructed velocity distribution reflects the stress conditions disturbed by mining activities. Examining Fig. 8 one can clearly distinguish the main regions with high and low pressure.

4. Conclusions

We can conclude that the proposed simulated annealing approach to seismic traveltime tomography is a very robust and stable method. It is shown that unlike local methods, the convergence of the simulated annealing algorithm is independent of the initial model: even in cases of virtually no prior information, when local optimization techniques fail, the simulated annealing procedure yields interpretable solutions. It can provide a number of acceptable solutions. When prior information is sparse, the solution of the annealing scheme can be used as an input to a local optimization technique producing an even more accurate result. It is noteworthy that simulated annealing, as any other global optimization methods, can be used for solving almost any types of inversion problems. In the near future we intend to extend the method to three dimensions and to test it in earthquake tomography and hypocenter location.

Acknowledgements

The reported investigation was financially supported by the Hungarian Scientific Research Fund (No. F019277). The author is also grateful to the Department of Mining Geophysics, ELGI for providing the field data in the study and would like to thank István Bondár for the many helpful and stimulating discussions. I am also thankful to M. Körnig and one anonymous reviewer for making various suggestions which have helped me in improving this paper.

References

Bohachevsky, I.O., Johnson, M.E., Stein, M.L., 1986. Generalized simulated annealing for function optimization. Technometrics 28, 209–217.

Dines, K.A., Lytle, R.J., 1979. Computerized geophysical tomography. Proc. IEEE 67, 1065–1073.

Gordon, R., 1974. A tutorial on ART. IEEE Trans. Nucl. Sci. NS 21, 78–93.

Herman, G.T., Lent, A., Rowland, S.W., 1973. ART: Mathematics and applications: a report on the mathematical foundations and on the applicability to real data of the algebraic reconstruction techniques. J. Theor. Biol. 42, 1–32.

Kirkpatrick, S., Gelatt, C.D. Jr., Vecchi, M.P., 1983. Optimization by simulated annealing. Science 220, 671–680.

Landa, E., Beydoun, W., Tarantola, A., 1989. Reference velocity model estimation from prestack waveforms: coherency optimization by simulated annealing. Geophysics 54, 984–990.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092.

Mosegaard, K., Vestergaard, P.D., 1991. A simulated annealing approach to seismic model optimization with sparse prior information. Geophys. Prospect. 39, 599–611.

Moser, T.J., 1991. Shortest path calculation of seismic rays. Geophysics 56, 59–67.

Peterson, J.E., Paulsson, B.N.P., McEvilly, T.V., 1985. Applications of algebraic reconstruction techniques to crosshole seismic data. Geophysics 50, 1566–1580.

Phillips, W.S., Fehler, M.C., 1991. Traveltime tomography: a comparison of popular methods. Geophysics 56, 1639–1649.

Pullammanappallil, S.K., Louie, J.N., 1994. A generalized simulated annealing optimization for inversion of first-arrival times. Bull. Seismol. Soc. Am. 84, 1397–1409.

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

Rothman, D.H., 1986. Automatic estimation of large residual statics corrections. Geophysics 51, 332–346.

Sen, M.K., Stoffa, P.L., 1991. Nonlinear one-dimensional seismic waveform inversion using simulated annealing. Geophysics 56, 1624–1638.

van der Sluis, A., van der Vorst, H.A., 1987. Numerical solution of large, sparse linear algebraic systems arising from tomographic problems. In: Nolet, G. (Ed.), Seismic Tomography. Reidel, Dordrecht, pp. 49–83.

van Laarhoven, P.J.M., Aarts, E.H.L., 1987. Simulated Annealing: Theory and Application. Reidel, Dordrecht.

Vasudevan, K., Wilson, W.G., Laidlaw, W.G., 1991. Simulated annealing static computation using an order-based energy function. Geophysics 56, 1831–1839.

Vestergaard, P.D., Mosegaard, K., 1991. Inversion of post-stack seismic data using simulated annealing. Geophys. Prospect. 39, 613–624.

Wéber, Z., 1995. Some improvement of the shortest path ray tracing algorithm. In: Diachok, O., Caiti, A., Gerstoft P., Schmidt, H. (Eds.), Full Field Inversion Methods in Ocean and Seismo-Acoustics. Kluwer Academic Publishers, Dordrecht, pp. 51–56.

White, S.R., 1984. Concepts of scale in simulated annealing. Proc. IEEE Int. Conference on Computer Design, Port Chester, pp. 646–651.