*Geophys. J. Int.* (1991) **106**, 379–385

# Magnetotelluric appraisal using simulated annealing

S. E. Dosso* and D. W. Oldenburg

*Department of Geophysics and Astronomy, University of British Columbia, 129–2219 Main Mall, Vancouver, BC, V6T 1W5, Canada*

Accepted 1991 February 21. Received 1990 December 10; in original form 1990 August 17

## SUMMARY

Conductivity models derived from magnetotelluric measurements can be appraised by constructing extremal models which minimize and maximize localized conductivity averages. These extremal models provide lower and upper bounds for the conductivity average over the region of interest. Previous applications of this method have constructed extremal models via (iterated) linearized inversion; however, it is difficult to verify that the computed bounds represent global (rather than local) extrema. In this paper, a method of constructing extremal models using simulated annealing optimization is developed. Simulated annealing requires no approximations and is renowned for its ability to avoid unfavourable local minima. The optimization procedure is flexible and general, and can be applied to construct models which extremize a linear or non-linear objective function in any inverse problem for which the corresponding forward solution exists. Appraisal via simulated annealing is demonstrated using synthetic data and field measurements, and the results are compared with those based on linearization. The comparisons suggest that the bounds calculated via linearization represent excellent approximations to the global extrema.

**Key words:** appraisal, extremal models, magnetotellurics, simulated annealing.

## 1 INTRODUCTION

The magnetotelluric (MT) method uses surface measurements of natural electromagnetic fields to investigate the subsurface conductivity distribution of the Earth. Determining useful information about the conductivity from a set of measured responses defines the MT inverse problem. The most common approach used to solve the inverse problem is to construct a model which adequately reproduces the data. Unfortunately, this inverse construction problem cannot be solved unambiguously: it can be shown that if one model exists which fits the data, then infinitely many such models exist. Even though it is possible to construct models of specific character or models which are close to an assumed model, a constructed solution provides no indication about the range of acceptable models.

Another approach to the inverse problem is that of appraisal. Rather than constructing one or more of the infinite number of possible model solutions, the goal of appraisal is to calculate properties which all acceptable models (including the true model) share. Backus & Gilbert (1968, 1970) developed a method of appraisal for linear inverse problems based on generating unique averages of the model from linear combinations of the data. This method can be applied to non-linear inverse problems such as the MT problem by linearizing about a reference model. Unfortunately, in this case the unique averages computed pertain only to models that are linearly close to the reference model. Oldenburg (1979) constructed a number of different conductivity models which fit a set of MT data, and found different values for the model average by linearizing about these models. Parker (1983) and Oldenburg, Whittall & Parker (1984) have found linearized Backus–Gilbert appraisal to be inadequate for the MT problem.

An alternative form of appraisal was presented by Oldenburg (1983, hereafter O1), who determined bounds for conductivity averages by constructing extremal models which minimize or maximize the conductivity over a specified region. In extremal model appraisal, the conductivity model is represented by a set of parameters $\{\sigma_1, \sigma_2, \ldots, \sigma_M\}$ which form a discretized representation of the function $\sigma(z)$ on a depth partition $\{0, z_1, z_2, \ldots, z_M\}$. Model limits $\sigma_i^- \leq \sigma_i \leq \sigma_i^+$ are required for each conductivity element. These limits represent the *a priori* knowledge about the model: if confining limits are known for some depths, then including this information can lead to improved computed bounds; alternatively, if reliable constraints are not available, then suitably wide limits may be assumed so as to not influence the solution (O1).

* Present address: Defence Research Establishment Pacific FMO, Victoria, BC, V0S 1B0, Canada.

A boxcar average of the conductivity over a width $\Delta$ centred at a depth $z_0$ can be represented as a linear combination of the model parameters

$$\bar{\sigma}(z_0, \Delta) = \sum_{i=1}^{M} w_i \sigma_i, \tag{1}$$

where

$$w_i = \begin{cases} (z_i - z_{i-1})/\Delta, & \text{if } z_0 - \Delta/2 \le z_{i-1},\ z_i \le z_0 + \Delta/2, \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The goal is to construct a model which minimizes or maximizes $\bar{\sigma}(z_0, \Delta)$ subject to achieving an acceptable fit to the data. To accomplish this, O1 linearized the relationship between the response functional $R$ and model $\sigma(z)$ by neglecting second-order terms in the expansion of $R$ about a starting model $\sigma_0(z)$ to obtain

$$\delta R_j = \int_0^\infty G_j(\sigma_0, z)\, \delta\sigma(z)\, dz, \qquad j = 1, \ldots, N, \tag{3}$$

where $\delta R_j = R_j - R_j(\sigma_0)$ is the difference between measured and predicted responses, $\delta\sigma$ represents a perturbation to the starting model, and $G_j(\sigma_0)$ is the Fréchet kernel. Substituting $\delta\sigma = \sigma - \sigma_0$ leads to

$$\delta R_j + \int_0^\infty G_j(\sigma_0, z)\sigma_0(z)\, dz = \int_0^\infty G_j(\sigma_0, z)\sigma(z)\, dz, \qquad j = 1, \ldots, N, \tag{4}$$

which represents a set of linear equations that relate (modified) responses directly to the model (rather than the model perturbation). These equations may be expressed in discretized form as

$$\tilde{R}_j = \sum_{i=1}^{M} \Gamma_{ji}\sigma_i \qquad j = 1, \ldots, N, \tag{5}$$

where $\tilde{R}_j$ represents the modified responses and $\Gamma_{ji} = \int_{z_{i-1}}^{z_i} G_j(z)\, dz$. Linear programming techniques may be used to minimize or maximize $\bar{\sigma}(z_0, \Delta)$ given by (1) subject to the constraint that the data equations (5) are satisfied according to the $\chi^1$ measure of misfit

$$\chi^1(\sigma) = \sum_{j=1}^{N} \left| \frac{R_j - R_j(\sigma)}{s_j} \right| = \chi^1_d, \tag{6}$$

where $s_j$ is the standard deviation of the $j$th datum. $\chi^1_d$ represents the desired level of misfit, which is generally taken to be $\sqrt{2/\pi}\,2N$, the expected value of the $\chi^1$ statistic for $N$ complex responses (Parker & McNutt 1980). Since higher order terms have been neglected, this procedure must be repeated iteratively until an acceptable model is achieved. If the construction procedure converges to the global extremum, then a true bound for the conductivity has been found. However, since the method uses linearization, there is always a danger that the computed bound may be representative of a local extremum. To investigate this possibility, O1 and Dosso & Oldenburg (1989) initiated their inversion algorithms with diverse starting models. Although the extremal models constructed sometimes differed in minor detail, they did not find a case where the computed model average differed; this provided confidence that the computed bounds were not dependent on the starting model.

To further investigate the extremal solution, this paper presents a new method of constructing extremal models for non-linear inverse problems which is not based on linearization. Rather, the construction problem is formulated as an optimization problem which is solved using the method of simulated annealing (Kirkpatrick, Gelatt & Vecchi 1983). Simulated annealing is an optimization procedure which has been successfully applied to many problems in the field of combinatorial optimization (see e.g. van Laarhoven & Aarts 1987). A major advantage of the method is its inherent ability to avoid unfavourable local minima. This feature is of crucial importance to the application here. Although appraisal using simulated annealing is considerably less efficient than linearized methods, it represents an independent method of estimating conductivity bounds and may be used to corroborate the results of the linearized appraisal. In addition, the simulated annealing approach is general and flexible, and can be applied to extremize any functional in any inverse problem for which the corresponding forward solution exists.

In the next section of this paper, the method of simulated annealing and its analogy with statistical mechanics is briefly presented [for a comprehensive treatment, see Kirkpatrick *et al.* (1983) and Kirkpatrick (1984), or the monograph by van Laarhoven & Aarts (1987)]. In Section 3 the simulated annealing appraisal algorithm is described, and in Section 4 examples of the analysis are presented and compared with results of the linearized appraisal for synthetic and field MT measurements.

## 2 SIMULATED ANNEALING

Simulated annealing is a general optimization procedure which mimics the thermodynamical process of annealing. Annealing is the process by which crystals are grown: a substance is first heated to melting, then cooled slowly until a crystal is formed. Simulated annealing draws an analogy between the parameters of an optimization problem and particles in an idealized physical system. The optimization procedure involves simulating the evolution of the physical system as it cools and anneals into a state of minimum energy.

Let possible configurations of the physical system be defined by a set of $M$ parameters $\mathbf{r} = \{r_1, r_2, \ldots, r_M\}$ which may represent, for example, particle positions and velocities. A fundamental result of statistical mechanics is the Boltzmann probability distribution

$$P(\mathbf{r}) \propto e^{-E(\mathbf{r})/kT}, \tag{7}$$

which relates the probability $P$ of the system at (absolute) temperature $T$ being in configuration $\mathbf{r}$ to the energy $E(\mathbf{r})$ ($k$ is Boltzmann's constant). According to the Boltzmann distribution, the probability function for a system in equilibrium at (non-zero) temperature is distributed over all possible configurations $\mathbf{r}$; thus, even at low temperature there is a finite chance of the system being in a high-energy configuration. At non-zero temperatures the system is continuously perturbed due to thermal agitation. According to (7), perturbations that increase the energy are allowed,

although they are less probable than perturbations that decrease the energy. As $T$ decreases, the Boltzmann distribution assigns progressively greater probability to low-energy configurations, and perturbations which increase $E$ become increasingly less likely. In the limit as $T \rightarrow 0$, the Boltzmann distribution collapses into the ground state for the system.

The ground state often corresponds to a pure crystal which represents the global minimum-energy configuration for the system. To achieve the ground state the system must be cooled very slowly to maintain an equilibrium distribution. If the system is allowed to get out of equilibrium it will not obtain the ground state, but rather forms a polycrystalline or amorphous (glass) state with no crystalline order and only locally optimal structure. These configurations represent local minima for the system energy.

A simple algorithm which simulates the average behaviour of a system of particles in thermal equilibrium was developed by Metropolis *et al.* (1953). In each step of the algorithm, a particle is given a small random displacement and the resulting change in the energy of the system $\Delta E$ is computed. The probability of such a change occurring is assumed to be

$$P(\Delta E) = e^{-\Delta E/kT}. \tag{8}$$

If $\Delta E < 0$ (i.e. the transition has lowered the system energy) the probability according to (8) is greater than unity; in this case the change is arbitrarily assigned a probability $P = 1$ and the transition is accepted. The case $\Delta E \geq 0$ is treated probabilistically: a random number $\xi$ is generated from a uniform distribution on the interval $[0, 1]$; if $\xi \leq P(\Delta E)$, the new configuration is retained; if not, the original configuration is retained. Repeating this basic step many times simulates thermal motion at a temperature $T$. The system eventually reaches equilibrium and the probability of a given configuration $\mathbf{r}$ evolves into a Boltzmann distribution.

Simulated annealing optimization is based on an analogy between the undetermined parameters of a mathematical system to be optimized and the particles of a physical system. The objective function of the optimization problem is analogous to the energy of the physical system, with the desired optimum corresponding to the ground-state configuration. Optimization via simulated annealing involves statistically modelling the evolution of the physical system using the Metropolis algorithm at a series of decreasing temperatures that allow the system to anneal into a state of minimum energy. Accepting perturbations to the system which increase the objective function as well as those which decrease it according to the Metropolis criterion allows the algorithm to escape from local minima.

In simulated annealing, the temperature $T$ acts as a control parameter and has the same units as the objective function $E$ ($k$ is taken to be 1). The process begins with the system in a known configuration and a procedure of generating random perturbations or changes in the configuration. The initial temperature must be high enough so that the system is completely 'melted', i.e. so that essentially all changes are accepted according to the Metropolis criterion (8) regardless of whether $E$ is decreases or increased. This completely disorders the system and renders the solution independent of the initial configuration.

The temperature is reduced in stages allowing enough perturbations at each temperature that the system reaches equilibrium. As $T$ is decreased, the probability of accepting configuration changes that increase the objective function decreases. Finally, at a low temperature the system 'freezes' and no further changes are accepted. The sequence of temperatures and number of perturbations at each temperature is referred to as the annealing schedule. An appropriate annealing schedule and effective method of perturbing the system is generally problem-specific and may require experimentation (Kirkpatrick *et al.* 1983).

The analogy between cooling a fluid and optimizing a function of many parameters may fail in one respect. In an ideal fluid the particles are all identical and the ground state is distinct and well defined, corresponding to a regular crystal lattice with a high degree of symmetry. However, mathematical optimization problems often involve many distinct, non-interchangeable elements which make a regular, symmetric solution unlikely (Kirkpatrick *et al.* 1983). Also, energy functions with terms representing conflicting objectives (and therefore favouring incompatible configurations) may preclude a simple well-ordered solution to the optimization problem (van Laarhoven & Aarts 1987). Optimizations with these characteristics are termed 'frustrated'. In physical examples of frustrated systems (e.g. magnetic 'spin glasses', see van Laarhoven & Aarts 1987) the low-temperature states are degenerate so that a number of near-ground-state configurations exist with essentially identical energy. Similarly, in optimization problems frustration makes the search for the global optimum more difficult. However, the degeneracy implies that there should be many equivalent solutions which closely approximate the absolute minimum. In practice, finding one of these solutions is sufficient (Kirkpatrick 1984).

## 3 THE SIMULATED ANNEALING APPRAISAL ALGORITHM

The construction of extremal models may be formulated as a simulated annealing optimization problem as follows. The system to be optimized is represented by $\boldsymbol{\sigma} = \{\sigma_i\}$ and the ensemble of possible system configurations is taken to be the set of all configurations $\{\sigma_i\}$, such that $\sigma_i^- \leq \sigma_i \leq \sigma_i^+$. Since each $\sigma_i$ is allowed to vary continuously between its limits, there are an infinite number of possible configurations, and therefore the formulation here is not strictly a combinatorial optimization problem. However, with an appropriate procedure of perturbing the system it is straightforward to apply simulated annealing to this problem. Vanderbilt & Louie (1984) present a method of applying simulated annealing to continuous problems when no limits can be estimated for the system parameters.

Extremal models which minimize the conductivity average $\bar{\sigma}(\Delta)$ may be constructed by minimizing the energy or objective function

$$E(\boldsymbol{\sigma}, \Delta) = |\chi^1(\boldsymbol{\sigma}) - \chi^1_d| + \beta \sum_{i=1}^{M} w_i \sigma_i. \tag{9}$$

The first term on the right of (9) represents the difference between the achieved and desired $\chi^1$ misfit, the second term represents the model average to be minimized, and $\beta$ is a

trade-off or scaling parameter which determines the relative importance of the misfit and model average in the minimization. To construct an extremal model which maximizes $\bar{\sigma}$, the objective function (9) is modified to be

$$E(\boldsymbol{\sigma}, \Delta) = |\chi^1(\boldsymbol{\sigma}) - \chi^1_d| + \beta \sum_{i=1}^{M} w_i(\sigma_i^+ - \sigma_i). \tag{10}$$

Requiring a precise level of fit to the data according to the $\chi^1$ criterion is adopted here for simulated annealing appraisal so that the results may be compared directly with those of the linearized (LP) appraisal algorithm. However, it should be noted that the simulated annealing approach is not restricted to this measure of misfit, and in other applications it may be appropriate to use the more conventional $\chi^2$ misfit measure (e.g. Parker 1977). Likewise, the simulated annealing optimization method can be adapted to minimize other functionals of the model (e.g. a measure of the model structure), subject to fitting the data, by replacing the second term of (9) or (10) by the functional to be minimized.

The energy functions (9) and (10) generally lead to frustrated optimization problems since the two terms represent conflicting objectives of fitting the data and minimizing or maximizing the model average. This implies that there are likely many solutions which closely approximate the optimal solution and are equally acceptable (Kirkpatrick 1984).

The basic step at each temperature of the annealing schedule involves perturbing the system $\boldsymbol{\sigma}$, computing the resulting change in the objective function $\Delta E$, and accepting or rejecting the new configuration based on the Metropolis criterion (8). System perturbations involve randomly changing one or more conductivity elements. A conductivity element $\sigma_i$ is changed according to

$$\sigma_i = \sigma_L + \eta(\sigma_U - \sigma_L), \tag{11}$$

where $\eta$ is a random number from a uniform distribution on [0, 1]. In (11), $\sigma_L$ and $\sigma_U$ are initially taken to be $\sigma_i^-$ and $\sigma_i^+$ so that $\sigma_i$ can take on any value between its limits. After a sufficient number of temperature steps, large-scale structure of the solution becomes (relatively) fixed and extreme perturbations will inevitably be rejected. At this point $\sigma_L$ and $\sigma_U$ may be reset to $\sigma_i/2$ and $2\sigma_i$ (if these values are within the limits). A perturbation can involve changing one element (chosen sequentially) or a random combination of elements. Combinations involve a random number of elements which are chosen at random and changed according to (11). We have found the most effective manner of perturbing the system involves alternating between changing a single element and a combination of elements, and cycling through the system a number of times. Typically, this would involve 300–500 perturbations per temperature.

Since finding the best possible extremal value for $\bar{\sigma}$ is crucial, we have adopted a cautious approach to the annealing schedule. An initial temperature $T_0$ is chosen so that at least 90 per cent of the perturbations are accepted. This effectively melts the starting model. The temperature is reduced according to the sequence

$$T_{i+1} = \varepsilon_T T_i, \qquad i = 0, 1, \ldots, \tag{12}$$

where $\varepsilon_T$ is typically 0.99. The temperature reduction continues until the system freezes and no further perturbations are accepted. Although faster annealing schedules could likely be devised by reducing $T$ more quickly at high and low temperatures, this schedule has proved very effective.

Determining appropriate values for the trade-off parameter $\beta$ is another important aspect of the optimization problem. Our optimization procedure is guided by the following ideas. The objective function (9) or (10) may be represented as

$$E(\boldsymbol{\sigma}, \Delta) = E_\chi + \beta E_{\bar{\sigma}}, \tag{13}$$

where $E_\chi$ is the misfit component, $E_{\bar{\sigma}}$ is the model average component, and $\beta$ is the trade-off parameter which controls the relative importance of the two contributions in the objective function. In the early stages of the annealing process the misfits are large and $\beta$ is chosen such that $\beta E_{\bar{\sigma}} > E_\chi$. This generally causes the conductivity to approach the imposed limits in the depth range over which the conductivity bound is desired. The data usually are not fit well at this stage. In order to lessen the tendency for the models to be close to the limits, the value of $\beta$ must be gradually reduced. This also allows models to be generated which greatly reduce $E_\chi$. At some point in the optimization the desired misfit is (approximately) achieved and $E_\chi \approx 0$. We wish to preserve that condition, and hence it is necessary that $\beta$ be sufficiently small so that even small deviations of $\chi^1$ from $\chi^1_d$ are discriminated against. When $\beta$ achieves such a value, it is not necessary to reduce it further. Effectively, only model perturbations which are consistent with keeping $E_\chi$ small have a chance of being accepted, and whether these perturbations are accepted then depends primarily upon how they affect $E_{\bar{\sigma}}$. This latter decision is unaffected by scaling $E_{\bar{\sigma}}$ by a constant. In summary, $\beta$ is initially chosen to be large. As the temperature is reduced in the annealing schedule $\beta$ is also reduced until it is sufficiently small, at which point it can be kept constant.

We have not investigated the ideal relationship between $E_\chi$, $E_{\bar{\sigma}}$, $\beta$ and $T$, but the ideas presented here indicate that $\beta$ should decrease with temperature. We have chosen an explicit relationship

$$\beta(T_{i+1}) = \varepsilon_\beta \beta(T_i), \qquad i = 0, 1, \ldots, \tag{14}$$

where $\varepsilon_\beta \leq 1$ (equality is invoked once model structure which acceptably fits the data becomes relatively permanent). Appropriate values of $\beta(T)$, as determined by $\beta(T_0)$ and $\varepsilon_\beta$ in (14), are generally problem-dependent and may require some experimentation. However, we have found that the procedure works well and that the extremal bound obtained for $\bar{\sigma}$ is independent of the precise values chosen, provided $\beta$ is varied in the manner described above.

## 4 APPRAISAL EXAMPLES

In this section simulated annealing appraisal is illustrated for synthetic and field MT measurements and the results are compared to those of linearized extremal model appraisal. The synthetic test case considered here was used by Whittall & Oldenburg (1991) in their survey of 1-D MT inversion techniques. The true model consists of four homogeneous layers overlying a uniform half-space, and is indicated by the

dashed line in Figs 1 and 2. Complex responses consisting of ratios of orthogonal components of magnetic and electric field $R = B/E$ were generated at 12 periods equally spaced in logarithmic time from 0.0025 to 250 s. Accurate data are used in the appraisal, however, an uncertainty of 1.8 per cent in all responses is assumed so that the $\chi^1$ statistic can be used to measure the relative fit of the models. The desired misfit was taken to be $\chi^1_d = 19 \approx \sqrt{2/\pi}\,2N$. The model partition consists of $M = 50$ elements with assumed conductivity limits of $0.002 \leq \sigma_i \leq 0.2\ \mathrm{S\,m^{-1}}$.

Figures 1(a) and (b) show the extremal models which maximize $\bar{\sigma}(z_0 = 1300,\ \Delta = 800)$ constructed using simulated annealing and the linearized algorithm of Dosso & Oldenburg (1989), respectively. The similarity of the solutions near the region of maximization (900–1700 m depth) is amazing considering the completely different approaches of the two methods. The deep structure of the extremal models in Figs 1(a) and (b) differs somewhat; this structure is well removed from the region of maximization and does not affect the extremal value for $\bar{\sigma}$. These differences may reflect the degeneracy of the near-optimal solutions. The upper bounds for $\bar{\sigma}$ computed from these models are essentially identical: 0.0760 (simulated annealing) and 0.0767 S m$^{-1}$ (linearized). Both constructed models achieve the desired misfit of $\chi^1 = 19.0$. The true data (squares with error bars) and predicted responses (solid line) are compared in the panels on the right of the models in terms of apparent conductivity $\sigma_a$ and phase $\phi$.

Figures 2(a) and (b) show the extremal models which minimize $\bar{\sigma}(z_0 = 1300,\ \Delta = 800)$ constructed using simulated annealing and linearized inversion, respectively. The solutions are very similar near the region of minimization and yield lower bounds for $\bar{\sigma}$ of 0.0022 S m$^{-1}$ (simulated annealing) and 0.0020 S m$^{-1}$ (linearized). These computed lower bounds reflect the *a priori* limit of $\sigma^- = 0.002\ \mathrm{S\,m^{-1}}$, which indicates that the MT data essentially can not impose a non-zero lower bound on the conductivity for this averaging width (O1). As $\Delta$ increases the ability to resolve $\bar{\sigma}$ improves. Both constructed models shown in Fig. 2 achieve the desired misfit of $\chi^1 = 19.0$.

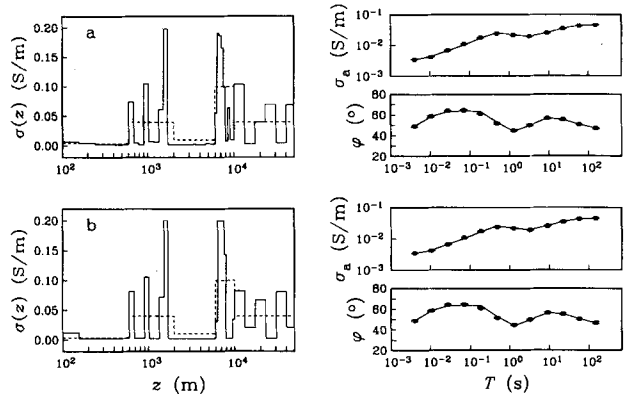


**Figure 1.** Extremal models constructed by maximizing $\bar{\sigma}(z_0 = 1300,\ \Delta = 800)$ for the synthetic MT example with model limits $0.002 \leq \sigma_i \leq 0.2\ \mathrm{S\,m^{-1}}$. (a) shows the model constructed using simulated annealing; (b) shows the solution from the linearized inversion. The two models are in excellent correspondence in and near the region where the conductivity is maximized. The true model is indicated by the dashed line. The constructed models have a misfit of $\chi^1 = 19.0$; the fit to the true responses is shown in the panels to the right.

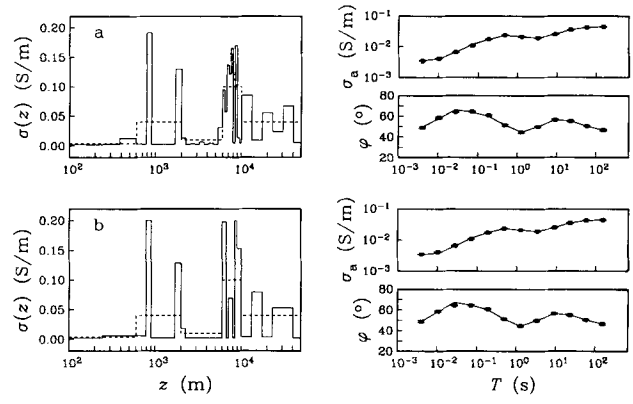


**Figure 2.** Extremal models constructed by minimizing $\bar{\sigma}(z_0 = 1300,\ \Delta = 800)$ for the synthetic MT example with model limits $0.002 \leq \sigma_i \leq 0.2\ \mathrm{S\,m^{-1}}$. (a) shows the model constructed using simulated annealing; (b) shows the solution from the linearized inversion. The two models are in excellent correspondence in and near the region where the conductivity is minimized. The true model is indicated by the dashed line. The constructed models have a misfit of $\chi^1 = 19.0$; the fit to the true responses is shown in the panels to the right.

It is interesting to compare the extremal models with the theoretical results of Weidelt (1985). Weidelt analytically treated the problem of extremizing the conductance function $S(z_2) = \int_0^{z_2} \sigma(z)\,dz$ subject to exactly fitting a small number of (precise) MT responses. Weidelt's solutions consist of thin zones of infinite conductivity, but finite conductance, embedded in an insulating half-space. When $S$ is maximized, a conductive zone is just included at the edge of the region of maximization; when $S$ is minimized, a conductive zone is just excluded at the edge of the region of minimization. The extremal models shown in Figs 1 and 2 resemble Weidelt's solutions with thin conducting zones just included or excluded at either edge of the region of maximization or minimization, respectively. This would seem to suggest that both the simulated annealing and linearized extremizations approach discretized approximations to the global extremal solution.

In O1, upper and lower bounds $\sigma^U(\Delta)$ and $\sigma^L(\Delta)$ were computed for $\bar{\sigma}(z_0, \Delta)$ at a number of different averaging widths $\Delta$, and plotted as a function of $\Delta$ to illustrate the resolution of the data at the depth $z_0$. Fig. 3 shows the

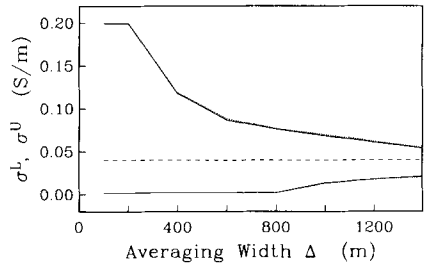


**Figure 3.** Computed lower and upper bounds for $\bar{\sigma}(z_0 = 1300,\ \Delta)$ for the synthetic MT example with model limits $0.002 \leq \sigma_i \leq 0.2\ \mathrm{S\,m^{-1}}$. The solid line indicates bounds from linearized appraisal, the dotted line indicates bounds from linearized appraisal, and the true model averages are indicated by the dashed line.

**Table 1.** Summary of lower and upper bounds computed for $\bar{\sigma}(z_0 = 1300, \Delta)$ using simulated annealing and linearized inversion to compute extremal models for the synthetic MT example.

| $\Delta$ (m) | $\bar{\sigma}$ (S/m) minimization | | $\bar{\sigma}$ (S/m) maximization | |
|---|---|---|---|---|
| | annealing | linearized | annealing | linearized |
| 100 | 0.0022 | 0.0020 | 0.200 | 0.200 |
| 200 | 0.0022 | 0.0020 | 0.200 | 0.200 |
| 400 | 0.0022 | 0.0020 | 0.118 | 0.119 |
| 600 | 0.0022 | 0.0020 | 0.0866 | 0.0884 |
| 800 | 0.0022 | 0.0020 | 0.0760 | 0.0767 |
| 1000 | 0.0134 | 0.0127 | 0.0680 | 0.0696 |
| 1200 | 0.0174 | 0.0172 | 0.0604 | 0.0616 |
| 1400 | 0.0211 | 0.0200 | 0.0537 | 0.0545 |

bounds computed by minimizing and maximizing $\bar{\sigma}(z_0 = 1300, \Delta)$ for eight averaging widths $\Delta$. The solid line indicates the bounds calculated using simulated annealing, the dotted line indicates the bounds from the linearized analysis and the dashed line shows the true model averages. The bounds computed using the two methods are almost indistinguishable over the entire range of $\Delta$, indicating that these methods produce virtually identical extremal values for $\bar{\sigma}$; these values are summarized in Table 1. It is noted, however, that the extremal values computed from the linearized analysis are slightly better at each value of $\Delta$ than those achieved by the simulated annealing method (i.e. the linearized method yields larger upper bounds and smaller lower bounds).

A final example of appraisal using simulated annealing considers a set of wide-band MT field data measured near Kootenay Lake in southeastern British Columbia, Canada, by Jones *et al.* (1988) as part of the LITHOPROBE Southern Cordilleran transect. Dosso & Oldenburg (1989) applied linearized extremal model appraisal to this data set to investigate an apparent low-conductivity region at 2000–7000 m depth indicated by minimum-structure models. Considering this data set of complex responses at $N = 34$ frequencies and a fine depth partition of $M = 130$ elements represents a demanding test for the simulated annealing appraisal algorithm. Model limits of $0.0001 \leq \sigma_i \leq 1.0\ \mathrm{S\,m^{-1}}$ are assumed, and the desired misfit was taken to be $\chi^1_d = 95$, since previous modelling studies indicated that the data could not be fit to within the expected value of $\chi^1$ (Dosso 1990). The upper bound for the model average over the low-conductivity region computed using simulated annealing is $0.0021\ \mathrm{S\,m^{-1}}$ and the constructed extremal model is shown in Fig. 4(a). By comparison, an upper bound of $0.0023\ \mathrm{S\,m^{-1}}$ was computed using the linearized algorithm, and the extremal models is shown in Fig. 4(b). Both constructed models have a misfit of $\chi^1 = 95.0$. The two models are in good agreement near the region of maximization 2000–7000 m depth. The conductivity remains near the lower limit in this region with narrow conductive zones just included at either edge of the region of maximization.

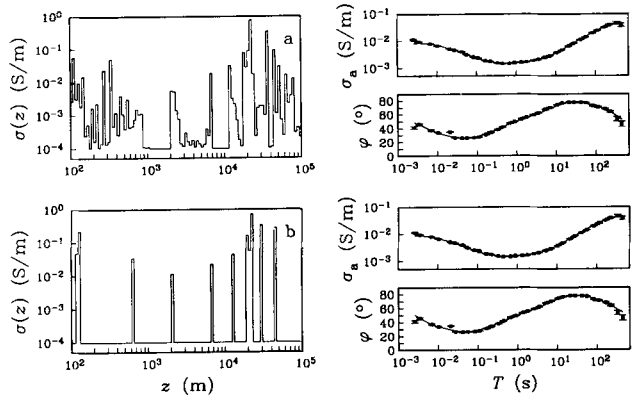


**Figure 4.** Extremal models constructed by maximizing $\bar{\sigma}$ over the apparent low conductivity region 2000–7000 m depth for the LITHOPROBE MT data set. Model limits are $0.0001 \leq \sigma_i \leq 1.0\ \mathrm{S\,m^{-1}}$. (a) shows the model constructed using simulated annealing; (b) shows the solution from the linearized inversion. The two models are in good correspondence in and near the region where the conductivity is maximized. The computed upper bounds for $\bar{\sigma}$ are $0.0021\ \mathrm{S\,m^{-1}}$ using simulated annealing, and $0.0023\ \mathrm{S\,m^{-1}}$ using linearized inversion. The constructed models have a misfit of $\chi^1 = 95.0$; the fit to the true responses is shown in the panels to the right.

## 4 DISCUSSION

The simulated annealing procedure developed in this paper has great flexibility and can be applied to minimize or maximize any (linear or non-linear) functional of the model and/or misfit in any inverse problem for which a solution to the forward problem exists. The annealing method generally requires a large number of solutions to the forward problem and therefore the computational efficiency depends directly on the efficiency of the forward solution.

Simulated annealing appraisal for the MT problem can be quite slow and is considerably less efficient than linearized methods. The solutions for the synthetic MT example required, on average, almost two days CPU time per extremization on a SUN 4/310 workstation, and represents a very careful approach to the annealing schedule to ensure that the best possible results are obtained. By comparison, the linearized extremizations required only about 3–5 min of computation time. However, since simulated annealing is well-known for its ability to avoid unfavourable local minima, it provides a useful method of corroborating the results of linearized analysis.

In principal, simulated annealing can find global minima (e.g. van Laarhoven & Aarts 1987) and therefore the bounds computed using the simulated annealing algorithm might be expected to be better than those of the linearized analysis. However, the success of simulated annealing optimization depends on the annealing schedule, the method of perturbing the system, and the choice of trade-off parameter $\beta$. We have investigated a number of possibilities for each, and have presented our best optimization algorithm. In all the cases considered, we have found that the extremal value for $\bar{\sigma}$ computed using simulated annealing is very close, but slightly inferior, to that computed using linearized appraisal. This indicates that the linearized approach produces excellent extremal values which in many cases may represent the best approximation (for a given depth partition) to the global extremum. The similarity in form of the constructed models to the theoretical extremal solutions would seem to support this conclusion. Consequently, even though the analysis in this

paper does not constitute a proof of global optimality, it does provide compelling evidence that linearized extremization is an efficient method of computing meaningful conductivity bounds.

## ACKNOWLEDGMENTS

This research was funded by NSERC grant 5-84270. SED was supported by an NSERC Postgraduate Scholarship.

## REFERENCES

Backus, G. E. & Gilbert, F., 1968. Numerical applications of a formalism for geophysical inverse problems, *Geophys. J.R. astr. Soc.*, **13,** 247–276.

Backus, G. E. & Gilbert, F., 1970. Uniqueness in the inversion of inaccurate gross earth data, *Phil. Trans. R. Soc. Lond.*, A, **266,** 123–192.

Dosso, S. E., 1990. Inversion and appraisal for the one-dimensional magentotellurics problem, *PhD thesis,* University of British Columbia, Vancouver.

Dosso, S. E. & Oldenburg, D. W., 1989. Linear and non-linear appraisal using extremal models of bounded variation, *Geophys. J. Int.*, **99,** 483–495.

Jones, A. G., Kurtz, R. D., Oldenburg, D. W., Boerner, D. E. & Ellis, R., 1988. Magnetotelluric observations along the LITHOPROBE Canadian Cordilleran Transect, *Geophys. Res. Lett.*, **15,** 677–680.

Kirkpatrick, S., 1984. Optimization by simulated annealing: quantitative studies, *J. Stat. Phys.*, **34,** 975–986.

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

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

Oldenburg, D. W., 1979. One-dimensional inversion of natural source magnetotelluric observations, *Geophysics,* **44,** 1218–1244.

Oldenburg, D. W., 1983. Funnel functions in linear and nonlinear appraisal, *J. geophys. Res.*, **88,** 7387–7398.

Oldenburg, D. W., Whittall, K. P. & Parker, R. L., 1984. Inversion of ocean bottom magnetotelluric data revisted, *J. geophys. Res.*, **89,** 1829–1833.

Parker, R. L., 1977. Understanding inverse theory, *Ann. Rev. Earth planet. Sci.*, **5,** 35–64.

Parker, R. L., 1983. The magnetotelluric inverse problem, *Geophys. Surv.*, **6,** 5–25.

Parker, R. L. & McNutt, M. K., 1980. Statics of the one-norm misfit error, *J. geophys. Res.*, **85,** 4429–4430.

Vanderbilt, D. & Louie, S. G., 1984. A Monte Carlo simulated annealing approach to optimization over continuous variables, *J. Comp. Phys.*, **36,** 259–271.

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

Weidelt, P., 1985. Construction of conductance bounds from magnetotelluric impedances, *J. Geophys.*, **57,** 191–206.

Whittall, K. P. & Oldenburg, D. W., 1991. Inversion of magnetotelluric data over a one-dimensional earth, in *Magnetotellurics in Geophysical Exploration,* ed. Wannamaker, P. E., SEG publication, in press.