Geophys. J. Int. (1989) **99**, 483–495

Linear and non-linear appraisal using extremal models of bounded variation

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 1989 April 18. Received 1989 April 16; in original form 1988 June 29

SUMMARY

Model features may be appraised by computing upper and lower bounds for the average value of the model over a specified region. The bounds are computed by constructing extremal models which maximize and minimize this average. In order to compute the most meaningful bounds, it is important that the allowed models are geophysically realistic. In this paper, the appraisal analysis of Oldenburg (1983) is extended to incorporate a bound on the total variation of the extremal models. Restricting the variation discriminates against highly oscillatory models and, as a consequence, the difference between upper and lower bounds is often considerably reduced. The original presentation of the funnel function bound curves is extended to include the variation of the model as another dimension. The interpreter may make use of any knowledge or insight regarding the variation of the model to generate realistic extremal models and meaningful bounds.

The appraisal analysis is extended to non-linear problems by altering the usual linearized equations so that a global norm of the model can be used in the objective function. The method is general, but is applied here specifically to compute bounds for localized conductivity averages of the Earth by inverting magnetotelluric measurements. The variation bound may be formulated in terms of conductivity or log conductivity. The appraisal is illustrated using synthetic data and field measurements from southeastern British Columbia, Canada.

Bounding the total variation may be viewed as constraining the flatness of the model. This suggests a new method of calculating (piecewise-constant) l_1 flattest models by minimizing the norm of the total variation. Unlike l_2 flattest models which vary in a smooth, continuous manner, the l_1 minimum-variation model is a least-structure model that resembles a layered earth with structural variations occurring at distinct depths.

Key words: appraisal, extremal models, inference, inversion, non-uniqueness, variation.

1 INTRODUCTION

In linear inverse theory, the general relationship between the data $\{e_j, j = 1, N\}$ and the model $m(z)$ is given by a Fredholm equation of the first kind

$$e_j = (m, g_j) = \int_0^a m(z) g_j(z)\, dz \qquad j = 1, \ldots, N, \tag{1}$$

where $g_j(z)$ is the kernel function corresponding to the jth datum. The model m and the kernels $\{g_j\}$ are members of a Hilbert space H and are defined on the interval $[0, a]$. The notation (,) is used to indicate the inner product. The goal of inverse theory is to use the observed data, their assumed errors and the known kernel functions to extract information about the model. A fundamental difficulty is that of non-uniqueness: for any finite data set, if there exists one model that adequately reproduces the data via (1), then infinitely many such models exist. These acceptable models may be diverse. As a result of the inherent non-uniqueness, any finite data set can not impose bounds on the value of the model at a fixed point. However, model averages over a finite width will, in general, be constrained by the data, provided at least one kernel function is non-zero over a portion of this width.

One approach to overcoming the non-uniqueness has been given by Backus & Gilbert (1970). By taking

appropriate linear combinations of the data equations, they generated unique averages of the model at a depth of interest z_0 of the form

$$\langle m(z_0)\rangle = (m, A(z_0)),$$

where $A(z, z_0) = \sum_{j=1}^{N} \alpha_j(z_0) g_j(z)$ is known as the averaging function or resolving kernel, and the coefficients $\alpha_j(z_0)$ are chosen to make $A(z, z_0)$ close (in some sense) to a Dirac delta function centred at z_0. The model average $\langle m(z_0)\rangle$ is unique in the sense that the inner product of $A(z, z_0)$ with any acceptable model must produce this same value. In practical cases where the data are inaccurate, a trade-off exists between the resolution width of the averaging function and the variance of the model average. The interpreter must select an $A(z, z_0)$ and associated $\langle m(z_0)\rangle$ which represents the most meaningful compromise between resolution and accuracy. Although this produces excellent results in some cases, for certain problems the averaging function $A(z, z_0)$ may have undesirable characteristics such as significant sidelobes or negative values, or it may not be centred at the depth of interest. In such cases, the average value, although unique, is not readily interpreted. Huestis (1987) presented a method for computing non-negative averaging functions, however, he demonstrated that for some problems such functions do not exist, and in cases where they do exist, the advantage gained in their use may be offset by a greatly increased computational burden. Another difficulty with the general formulation is that all mathematically acceptable models are included in the Backus–Gilbert averages. It may be that better averages could be obtained if additional physical constraints (e.g. positivity) could be incorporated.

To overcome these shortcomings, it is advantageous to obtain quantitative information about model averages by formulating the appropriate inference problem. The mathematical foundation for inference theory has been presented by Backus (1970a,b,c, 1972) and a pragmatic implementation of aspects of that philosophy has been presented by Oldenburg (1983) and will be briefly recounted here. In Oldenburg (1983) (henceforth referred to as O1) it was shown that upper and lower bounds for predicted linear functionals of the model could be computed using linear programming (LP) techniques. One of the most useful linear functionals is the integral of the model with a unimodular boxcar B of width Δ centred at a particular depth of interest z_0:

$$B(z, z_0, \Delta) = \begin{cases} 1/\Delta, & \text{if } |z - z_0| \le \Delta/2; \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

The resultant inner product

$$\bar{m}(z_0, \Delta) = (m, B(z_0, \Delta)) \tag{3}$$

represents an average value of the model over a width Δ centred at z_0. Since $B(z, z_0, \Delta)$ can not generally be formed as a linear combination of the kernel functions, $\bar{m}(z_0, \Delta)$ can not be determined uniquely. However, lower and upper bounds

$$m^{\mathrm{L}}(z_0, \Delta) \le \bar{m}(z_0, \Delta) \le m^{\mathrm{U}}(z_0, \Delta) \tag{4}$$

can be obtained by constructing extremal models which minimize and maximize (3) subject to the data and other available constraints on the model.

Numerical implementation of this procedure requires a partitioning of the interval $[0, a]$ with boundaries at $\{0, z_1, z_2, \ldots, z_n = a\}$. The model is assumed to be constant on each partition element, i.e. $m(z) = m_i$ for $z_{i-1} \le z \le z_i$. Linear programming methods can be used to minimize or maximize an objective function of the form $\phi = \sum_{i=1}^{n} w_i m_i$ subject to equality or inequality constraints on linear combinations of the model parameters m_i. The $\{w_i\}$ are a set of arbitrary weights which may be chosen according to

$$w_i = \begin{cases} (z_i - z_{i-1})/\Delta, & \text{if } z_0 - \Delta/2 \le z_{i-1},\ z_i \le z_0 + \Delta/2; \\ 0, & \text{otherwise,} \end{cases} \tag{5}$$

so as to make the objective function represent a discretized form of the model average, i.e.

$$\phi = \sum_{i=1}^{n} w_i m_i = \bar{m}(z_0, \Delta). \tag{6}$$

Lower and upper bounds m^{L} and m^{U} for $\bar{m}(z_0, \Delta)$ are calculated by minimizing and maximizing ϕ subject to the data constraints of (1), which may be written in discretized form as

$$e_j = \sum_{i=1}^{n} \gamma_{ij} m_i \qquad j = 1, \ldots, N \tag{7}$$

where

$$\gamma_{ij} = \int_{z_{i-1}}^{z_i} g_j(z)\, dz \qquad i = 1, \ldots, n. \tag{8}$$

Data inaccuracies may be incorporated either as 'hard' bounds (strict inequalities), or as 'soft' bounds where a few data are allowed to have large misfits while the total misfit is kept within acceptable limits. Soft bound constraints require additional LP variables to represent the data misfits.

An important advantage of the LP method is that any additional physical information about the model can be incorporated in the inversion, as long as it can be formulated in terms of a linear constraint. For instance, *a priori* lower and upper limits for the model elements

$$m_i^- \le m_i \le m_i^+, \tag{9}$$

are easily included.

For each value of z_0, the bounds $m^{\mathrm{L}}(z_0, \Delta)$ and $m^{\mathrm{U}}(z_0, \Delta)$ may be calculated for a number of different averaging widths Δ, and plotted as a function of Δ to produce a funnel function diagram. Such a plot provides immediate insight into the resolving power of the data at the depth of interest z_0. The only loss in generality in this formulation is that caused by the partitioning and parametrization. This is not of practical significance, however, provided that the partition quantization is sufficiently small. Lang (1985) demonstrated that the exact problem can be approximated to arbitrary accuracy by a discretized problem given a small enough partition interval.

Unfortunately, in O1 it was found that the constructed extremal models often exhibit unacceptably large oscillations. When model limits are absent or large, the extremal models are characteristically sparse and spiky, consisting of isolated regions of large amplitude. If confining model limits are imposed the extremal models characteristically consist of a sequence of sections which alternate between the imposed limits and in some cases fluctuate rapidly between the limits

(see, e.g. fig. 7 in O1). Although these models represent mathematically acceptable solutions, they are generally not geophysically realistic. As a consequence, since the funnel function bounds are obtained from these extremal models, it is likely that the bounds found in O1 are pessimistic. It is anticipated that more meaningful bounds could be calculated if these highly variable models are purposely winnowed from the analysis.

In this paper we introduce the total variation as a measure of the amount of structure of a model and discriminate against highly oscillatory models by placing an upper limit on the variation of the extremal models. In Section 2 the total variation is defined and two methods for carrying out the calculations are presented. In Section 3 the appraisal technique and the dependence of the computed bounds on the allowed variation is demonstrated using a simple linear example. Section 4 describes a general method by which the appraisal may be applied to non-linear problems, and in Sections 5 and 6 the method is applied specifically to the non-linear magnetotelluric problem by considering synthetic and field data cases.

2 EXTREMAL MODELS OF BOUNDED VARIATION

The total variation of a function $m(z)$ over an interval $[0, a]$ may be defined as (Korevaar 1968, p. 406)

$$V[m] = \int_0^a |dm|. \tag{10}$$

In order to eliminate extremal models which are judged to have too much structure, we will first supply an upper limit for V in (10), and then modify the method in O1 to include this constraint. By placing an upper bound on V, models which are sparse and spiky or jump repeatedly between the imposed limits can be discriminated against. Abrupt or discontinuous changes in the model are still allowed, but the total number and magnitude of such changes can be limited to an amount deemed reasonable. The goal is to select a variation bound which results in models that are judged to be geophysically realistic and produce the most meaningful funnel function bounds. Two methods of bounding the total variation of the extremal models are presented.

Method 1

The first method is applicable to models that are assumed to be continuous, i.e. $m \in C^1$. In this case (10) can be written as

$$V = V[m] = \int_0^a |m'(z)|\, dz,$$

where $m' = dm/dz$. In discrete form, where the model is assumed to have a constant gradient on each partition element, the variation can be written as

$$V = \sum_{i=1}^{n} |m_i'|(z_i - z_{i-1}). \tag{11}$$

Linear programming methods assume that the variables are positive, but model derivatives which may be either positive or negative can be accommodated by writing each parameter m_i' as the difference between two positive quantities $m_i' = p_i - q_i$, where $p_i, q_i \geq 0$ are the variables to be determined by the LP algorithm. The absolute values $|m_i'|$ in (11) can not be included in a linear constraint; however, they may be represented as $|m_i'| \leq p_i + q_i$, with equality holding when either p_i or q_i is zero. The total variation V must obey the inequality

$$V \leq \sum_{i=1}^{n} (p_i + q_i)(z_i - z_{i-1})$$

and an upper bound for the variation V_b may be specified by requiring

$$\sum_{i=1}^{n} (p_i + q_i)(z_i - z_{i-1}) \leq V_b. \tag{12}$$

Equation (12) is a linear constraint for the total variation of the form that can be included in the LP algorithm. To constrain the total variation according to (12), we need to write the LP objective function and constraints in a form which involves only p_i and q_i with $m_i' = p_i - q_i$. We first consider the objective function. If $m_0 = m(z = 0)$ is assumed known, then the value of the model on the ith partition is

$$m_i = m_0 + \sum_{k=1}^{i-1} (z_k - z_{k-1})m_k' + \tfrac{1}{2}(z_i - z_{i-1})m_i'. \tag{13}$$

The objective function in (6) becomes

$$\phi = m_0 + \sum_{i=1}^{n} (z_i - z_{i-1})\left(\tfrac{1}{2}w_i + \sum_{k=i+1}^{n} w_k\right)m_i'. \tag{14}$$

To put the data constraints into a compatible form, we integrate (1) by parts to obtain

$$f_j = \int_0^a m'(z)[h_j(z) - h_j(a)]\, dz, \tag{15}$$

where f_j and $h_j(z)$ are 'new' data and kernels given by

$$f_j = m_0 h_j(a) - e_j$$

$$h_j(z) = \int_0^z g_j(u)\, du.$$

Discretization yields

$$\begin{aligned} f_j &= \sum_{i=1}^{n} \gamma_{ij} m_i' \qquad j = 1, \ldots, N \\ \gamma_{ij} &= \int_{z_{i-1}}^{z_i} [h_j(z) - h_j(a)]\, dz \qquad i = 1, \ldots, n. \end{aligned} \tag{16}$$

As a final constraint, limits for individual model elements, $m_i^- \leq m_i \leq m_i^+$, may be included as

$$m_i^- \leq m_0 + \sum_{k=1}^{i-1} (z_k - z_{k-1})m_k' + \tfrac{1}{2}(z_i - z_{i-1})m_i' \leq m_i^+ \qquad i = 1, \ldots, n. \tag{17}$$

The LP problem of computing bounds for $\bar{m}(z_0, \Delta)$ consists of extremizing the objective ϕ given by (14) subject to the data constraints of (16), the model limits of (17) and the variation bound of (12). The extremal model may be calculated according to (13).

Method 2

The second method does not require the model to be a continuous function. The model is represented by a constant

value on each partition element. In this case the total variation of the model can be characterized as

$$V = \sum_{i=1}^{n-1} |m_{i+1} - m_i|. \tag{18}$$

Instead of formulating a LP problem in which the objective function, data and physical constraints, and model variation are all written in terms of the model derivative, we keep the formulation as presented in O1 and introduce $(n-1)$ additional variables

$$\Delta m_i = m_{i+1} - m_i \qquad i = 1, \ldots, n-1. \tag{19}$$

The variables to be solved for are the values of the model within the partition elements m_i and the change in the model at the partition boundaries Δm_i. These are not independent of course, and the $(n-1)$ equations (19) must be included as constraints. Unless the models are known to be monotonic, the Δm_i will usually be allowed to take on positive or negative values; this is accomplished by writing $\Delta m_i = r_i - s_i$, where $r_i, s_i \geq 0$. It follows that $|\Delta m_i| \leq r_i + s_i$ and therefore the total variation can be constrained by requiring

$$\sum_{i=1}^{n-1} (r_i + s_i) \leq V_b. \tag{20}$$

For the work presented here, we have programmed both methods. The advantage of the first method is that fewer variables and constraints are required in the LP algorithm. In Method 1, $2n$ variables are needed to represent the model derivative elements and the variation bound is specified as a single constraint, whereas in Method 2, $4n-2$ variables [n bipolar model elements and $2(n-1)$ variables to represent the model changes] and a total of n constraints are required to specify the variation bound. It would seem that Method 1 should be used in general. However, the sparsity of the constraint matrix is destroyed when the upper and lower limits for the model are expressed in terms of the model derivative. In practice, this can be a significant disadvantage since many LP algorithms are designed for large, sparse constraint matrices. Despite the fact that the second method requires more variables and constraints, the constraints are sparse, and for the LP code available to us, the second method was found both to be significantly faster and more stable computationally for large extremization problems. For this reason, and also because Method 2 does not require integrating the data equations or knowledge of any model value, the numerical examples given in this paper are computed using Method 2.

In either method, however, since the variation bound is specified as an inequality constraint, it may be that the extremal model does not achieve a total variation of V_b. The actual variation V of the constructed model is evaluated directly. In practice it is generally found that $V = V_b$, provided the variation bound V_b is less than the variation of the unbound extremal model.

The variation measure given by (11) or (18) is essentially the l_1 norm of the gradient energy of the model. The model which minimizes this norm is often referred to as the l_1 'flattest' or 'minimum-structure' model. Thus, bounding the total variation may also be viewed as constraining the flatness of the model. Any attempt to constrain the variation to a value less than that of the flattest model will result in an inconsistency between the variation bound and the data constraints.

This suggests a new method for constructing l_1 flattest models. The standard procedure (e.g. Oldenburg 1984) requires that the data equations be integrated by parts to produce data constraints in terms of the model derivative. Linear programming methods are used to construct the smallest derivative model, and this model is integrated (assuming a model endpoint value is known) to produce the flattest model. An alternate method would be to construct a piecewise-constant l_1 flattest model by formulating the LP problem in terms of model values and differences as variables and minimize the total variation $\phi = \sum_{i=1}^{n-1} |\Delta m_i|$. In practice, this is accomplished by minimizing

$$\phi = \sum_{i=1}^{n-1} (r_i + s_i). \tag{21}$$

This produces an l_1 flattest or minimum-variation model without resorting to model derivatives, and has the additional benefits of not requiring any integrations which may introduce numerical error, and not requiring knowledge of a model endpoint value.

Many minimum-structure inversion algorithms minimize the l_2 norm of the model gradient (e.g. Constable, Parker & Constable 1987; Smith & Booker 1988). This discriminates against large abrupt changes and the models produced tend to vary in a smooth, continuous manner with depth. Minimizing the l_1 total-variation norm given by (21) does not discriminate against abrupt changes, but rather produces a minimum-structure model that more closely resembles a layered earth with structural variations occurring at distinct depths.

3 LINEAR EXAMPLE

To illustrate the appraisal technique and to demonstrate the improvement in resolution that can result when a reasonable bound is placed on the total variation, consider the following numerical example used in O1. Let the model be defined on the interval [0, 1] as

$$m(z) = 1 - \tfrac{1}{2}\cos(2\pi z)$$

and the data be obtained from the equations

$$e_j = \int_0^1 m(z) e^{-(j-1)z}\, dz \qquad j = 1, \ldots, N.$$

A total of 11 data were generated, and these are used to infer information about the value of the true model for a depth $z_0 = 0.5$ where the model attains its maximum value of 1.5. Fig. 1(a) shows upper and lower bounds calculated using the method in O1 when no limits (except a positivity constraint) are placed on the model elements, and no bound is placed on the total variation. Averages of the true model are indicated by the dashed line. The wide bounds indicate that the resolving power of the data is poor. For instance, for an averaging width of $\Delta = 0.2$, the model average is known to lie only within the bounds $0 \leq \bar{m}(z_0, \Delta) \leq 4.16$ while the true model lies within the range $1.47 \leq m(z) \leq 1.50$. Only for $\Delta > 0.5$ is $m^L > 0$, so without additional

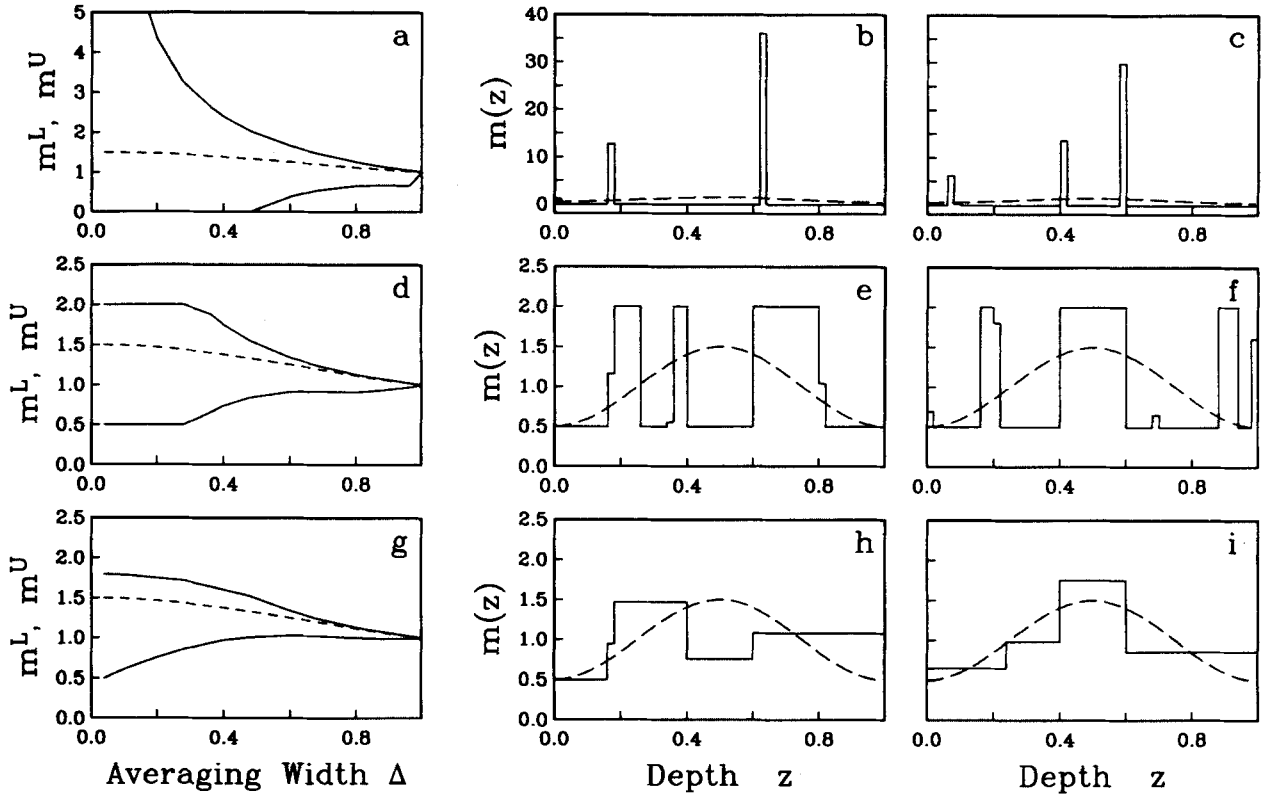

Figure 1. Lower and upper bounds for $\bar{m}(z_0 = 0.5, \Delta)$ are shown in (a), (d) and (g). In (a) only positivity was required, in (d) model limits $0.5 \leq m_i \leq 2.0$ have been imposed and in (g) a variation bound of $V_b = 2.0$ was also included. This variation bound corresponds to the actual variation of the true model. The true model averages are indicated by the dashed line. The two plots to the right of each funnel function diagram show the constructed models which minimize and maximize $\bar{m}(z_0 = 0.5, \Delta = 0.2)$. In these plots the true model is indicated by the dashed line.

physical information, a region of nonzero amplitude near $z_0 = 0.5$ is not a required feature of the model.

Figure 1(b and c) shows constructed extremal models which minimize and maximize the model average $\bar{m}(z_0, \Delta)$ for an averaging width of $\Delta = 0.2$. The true model is indicated by the dashed line. The constructed models consist of a sequence of zones of zero amplitude with two or three isolated regions of large amplitude extending over a width of one partition element. This structure is characteristic of all extremal models which produced the funnel functions of Fig. 1(a).

The constructed models for the unconstrained extremizations exhibit narrow regions with amplitudes of up to 40 or more. These values differ significantly from the true model. If reasonable limits for the model amplitude are known, the computed bounds can be greatly improved. Fig. 1(d) shows the funnel functions calculated after requiring that $0.5 \leq m_i \leq 2.0$. The significant improvement in resolution for all averaging widths is apparent when Fig. 1(a and d) are compared. The funnel functions also show the minimum resolution width required before the measured data influence the bounds. For instance, only for $\Delta > 0.28$ is $m^L > 0.5$, the imposed lower limit, and $m^U < 2.0$, the imposed upper limit.

Figure 1(e and f) shows constructed extremal models which minimize and maximize $\bar{m}(z_0, \Delta = 0.2)$. These models consist predominately of a sequence of sections which alternate between the imposed limits. Only a few model elements do not achieve either $m^- = 0.5$ or $m^+ = 2.0$. In some cases, the extremal models fluctuate rapidly between the imposed limits. An example of this is given in Fig. 2 which shows the constructed model which minimizes $\bar{m}(z_0, \Delta = 1.0)$.

The bimodal form of these extremizing models is similar to that of Parker's (1974, 1975) ideal bodies. The ideal body $m_I(z)$ is that model which is everywhere equal to either zero or M_0, where M_0 represents the greatest lower bound on the largest value of m (i.e. the smallest supremum of m). m_I is unique in that it is the only acceptable model which nowhere

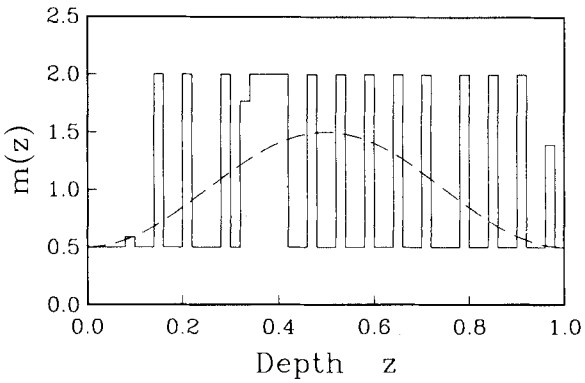

Figure 2. The constructed model which minimizes $\bar{m}(z_0 = 0.5, \Delta = 1.0)$ with imposed model limits $0.5 \leq m_i \leq 2.0$. The true model is indicated by the dashed line.

exceeds M_0. In the limit of $m^- \to 0$ and $m^+ \to M_0$, the LP extremal models will be equal to $m_1(z)$ regardless of the values of z_0 or Δ. However, we have found that the extremal models retain this bimodal form for a wide range of values for m^- and m^+, provided the discretization interval is sufficiently small.

Models such as those shown in Figs 1(b, c, e and f) and 2 might not be considered geophysically realistic, and hence the computed bounds may be unduly pessimistic. Fig. 1(g) shows the results of employing a variation bound $V_b = 2.0$, which represents the actual variation of the true model. A significant improvement in resolution is apparent when Figs 1(d) and (g) are compared. In this case there appears to be no minimum resolution width before the data influence the computed bounds. For instance, for an averaging width of $\Delta = 0.28$, the computed bounds are $0.86 \le \bar{m} \le 1.72$, while in Fig. 1(d), the computed bounds simply reflect the imposed limits $0.5 \le \bar{m} \le 2.0$. Fig. 1(h and i) shows constructed models which minimize and maximize $\bar{m}(z_0, \Delta = 0.2)$ for a variation bound $V \le 2.0$. These models do not exhibit excessive oscillations and might be considered to be more geophysically realistic than those shown in Figs 1(b, c, e and f) or 2.

To quantify the improvement in the bounds that results when the allowed variation is changed from V_1 to V_2, we compute the 'per cent improvement', P,

$$P(z_0, V_1, V_2, \Delta) = \left[1 - \frac{m^U_{V_2}(z_0, \Delta) - m^L_{V_2}(z_0, \Delta)}{m^U_{V_1}(z_0, \Delta) - m^L_{V_1}(z_0, \Delta)}\right] \times 100 \text{ per cent.} \quad (22)$$

In (22) the subscripts V_1 and V_2 indicate the total variation allowed in the extremal models. The results for $V_1 = \infty$ (no variation bound) and $V_2 = 2.0$ are shown in Fig. 3. For most averaging widths the funnel function bounds are improved by 30–40 per cent.

By reformulating the appraisal method to bound the total variation of the model, the analysis has been extended to include the variation as another dimension. Upper and lower bounds may now be considered as a function of both averaging width and model variation. Fig. 4 shows the computed bounds as a function of the allowed variation for fixed $\Delta = 0.2$. No limits (except positivity) were placed on the model elements. The true model has a variation of $V = 2.0$ and an average value of $\bar{m}(z_0, \Delta = 0.2) = 1.47$; this point is indicated by a cross. For large allowed variations the bounds are wide and the model average is poorly constrained. For instance, for a variation of $V = 14.0$, the model average is only known to lie within the bounds $0 \le \bar{m}(z_0, \Delta) \le 4.15$. As the allowed variation is decreased, the bounds converge smoothly. The upper bound decreases monotonically as the allowed variation is decreased from $V = 14.0$; however, it is not until the variation is less than about $V = 5.0$ that the lower bound increases from zero. At the true model variation of $V = 2.0$, the model average is known to lie within the bounds $0.73 \le \bar{m}(z_0, \Delta) \le 1.75$.

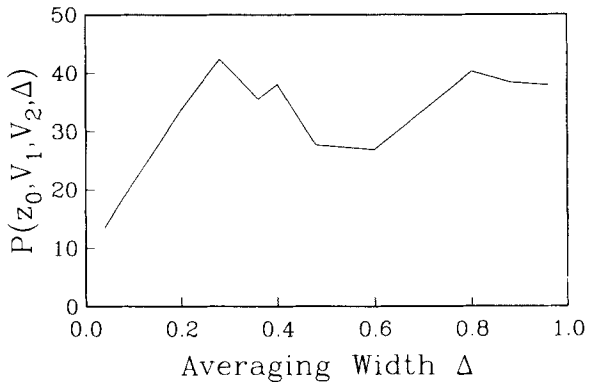

Figure 3. The per cent improvement, P, for $V_1 = \infty$ (no variation bound) and $V_2 = 2.0$, with model limits $0.5 \le m_i \le 2.0$.

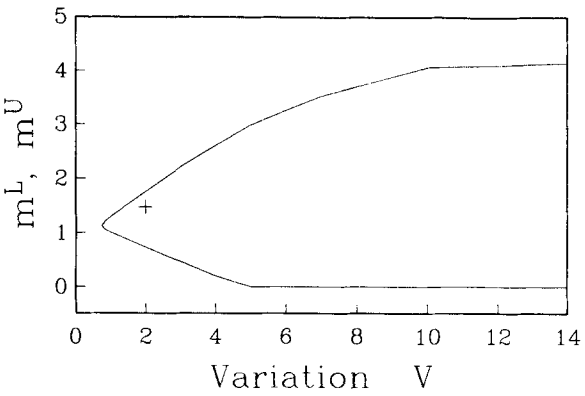

Figure 4. Lower and upper bounds as a function of the total variation V of the extremal model. The true model average is indicated by a cross.

Reducing the allowed variation to a value less than 2.0 excludes the true model from the LP solution space and may result in computed bounds which do not contain the true model average. This point is also illustrated in Fig. 4. As V is decreased below 2.0, the bounds continue to converge; for variations $V < 1.5$ the bounds no longer contain the true model average. The upper and lower bounds meet at a variation of $V = 0.75$. For this variation, the model average is known precisely, since $m^L = 1.15 \le \bar{m} \le 1.15 = m^U$, but this value does not correspond to the true model average of $\bar{m} = 1.47$. This demonstrates that although it is important to use the best possible value for the variation bound, over-constraining the variation (or any other physical property) can lead to misleading results.

The point at which the bounds meet represent the smallest possible variation which still permits an acceptable model. Any attempt to reduce the allowed variation below this value results in an inconsistency between the variation bound and the data constraints. The model which achieves the minimum variation is the (piecewise-constant) l_1 flattest model. The model constructed by minimizing (21) and the extremal models computed for $V_b = 0.75$ are identical and are shown in Fig. 5.

The funnel function resolution depends upon both the averaging width Δ and the allowed variation V. This dependence is illustrated in Fig. 6 in which contours of the normalized bound width $2(m^U - m^L)/(m^U + m^L)$ are plotted. The bound width increases with increasing variation and decreases with increasing averaging widths. The best resolution occurs for large averaging widths and small allowed variation. In practice, a plot like Fig. 6 together with an examination of the extremal models constructed for various variation bounds should enable an interpreter to

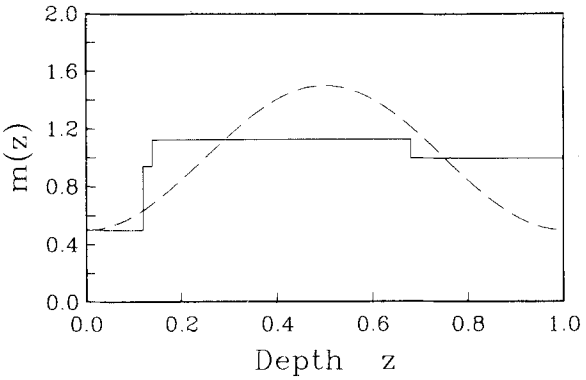

Figure 5. The l_1 minimum-variation model or piecewise-constant l_1 flattest model. The identical model is computed by minimizing the total-variation norm (21) or by minimizing or maximizing $\bar{m}(z_0, \Delta = 0.2)$ with $V_b = 0.75$. The true model is indicated by the dashed line.

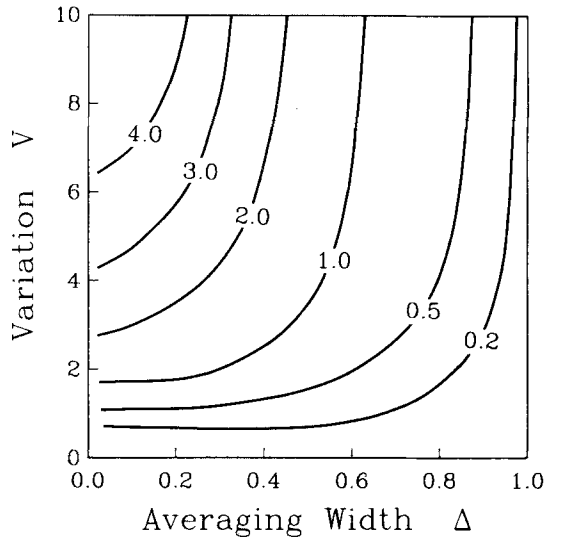

Figure 6. Contours of the normalized bound width $2(m^U - m^L)/(m^U + m^L)$ as a function of averaging width Δ and variation V for $z_0 = 0.5$.

determine meaningful bounds on the average value of the model over the region of interest.

4 APPRAISAL IN NON-LINEAR PROBLEMS

In many geophysical problems of interest the relationship between the model and the observations is functionally non-linear. When this is the case, Backus–Gilbert appraisal suffers a disadvantage in addition to those mentioned in Section I in that the unique averages computed pertain only to models that are linearly close to an acceptable constructed model. The method of appraisal using extremal models was extended to non-linear inverse problems in O1 by altering the usual linearized equations so that a global norm of the model could be used in the objective function. The method is general and can be applied to any non-linear problem. In this section the general formulation is briefly described; in Sections 5 and 6 the method is applied specifically to the non-linear magnetotelluric inverse problem.

In a non-linear inverse problem, the model is related to the observations by

$$e_j = F_j[m(z)] \qquad j = 1, \ldots, N,$$

where F_j is a non-linear functional. This general relationship, however, is not in a form which is amenable to our appraisal analysis. We proceed by writing the unknown model $m(z)$ as the sum of a known starting model $m_0(z)$ and an unknown perturbation $\delta m(z)$, i.e. $e_j = F_j[m_0(z) + \delta m(z)]$. Expanding the functional F_j about $m_0(z)$ leads to

$$e_j = F_j[m_0(z)] + \int_0^a G_j[m_0; z]\, \delta m(z)\, dz + \text{HOT}, \tag{23}$$

where HOT represents higher-order terms. If the higher-order terms of (23) can be shown to be of order $\|\delta m\|^2$, then F_j is Fréchet differentiable and G_j is called the Fréchet kernel. In this case the higher-order terms may be neglected resulting in a linear expression

$$\delta e_j = \int_0^a G_j[m_0; z]\, \delta m(z)\, dz, \tag{24}$$

where $\delta e_j = e_j - F_j[m_0(z)]$. The crucial step, as first described in O1, is to write $\delta m(z) = m(z) - m_0(z)$ so that (24) may be recast as

$$\delta e_j + \int_0^a G_j[m_0; z] m_0(z)\, dz = \int_0^a G_j[m_0; z] m(z)\, dz. \tag{25}$$

The left side of (25) consists of known quantities and may be considered modified data. By formulating the inversion problem according to (25), the methods of linear inverse theory may be used to construct a model $m(z)$ which extremizes a global norm of the model. Since higher-order terms have been neglected, (25) is not exact and the linearized inversion must be repeated iteratively until the constructed model adequately reproduces the observed data.

The linear form in (25) allows considerable flexibility: models of different global character can be constructed by the application of different norms to $m(z)$. For instance, globally flattest or smoothest models may be found by minimizing the l_2 norm of the first or second derivative of $m(z)$. These models are particularly useful in that they exhibit the minimum structure necessary to fit the data, and therefore are as simple as possible. Constable *et al.* (1987) and Smith & Booker (1988) have applied this method to the inversion of magnetotelluric data. In our minimum-structure inversion algorithm we minimize the norm

$$\int_0^a w(z) \left[\frac{dm(z)}{df(z)}\right]^2 df(z), \tag{26}$$

where $f(z)$ represents the depth function and $w(z)$ is an arbitrary weighting function. An alternative minimum-structure model is constructed by minimizing the total-variation norm given by (21). For the purposes of constructing extremal models of bounded variation to appraise model features, the norm to be extremized consists of the model average $\bar{m}$ given by (5) and (6).

5 APPLICATION TO MAGNETOTELLURICS

The goal of magnetotellurics (MT) is to infer information about the conductivity structure $\sigma(z)$ of the Earth by inverting electromagnetic responses observed on the Earth's surface. The method of appraisal using extremal models was applied to the non-linear MT problem in O1 to estimate upper and lower bounds for conductivity averages in regions of interest. However, the constructed extremal models were found to fluctuate in an unacceptable manner. To discriminate against such models and obtain more meaningful conductivity bounds, we have reformulated the method to include a bound on the total variation.

Magnetotelluric responses consist of ratios of magnetic and electric fields observed at discrete frequencies ω_j at the Earth's surface. We have chosen as data

$$e_j = \frac{B(\sigma; z=0, \omega_j)}{E(\sigma; z=0, \omega_j)} \qquad j = 1, \ldots, N,$$

where B and E represent orthogonal components of the magnetic and electric fields, as it appears to be the response for which the inverse problem is most linear (Smith & Booker 1988). When the complex response is considered separately as amplitude $|e_j|$ and phase ϕ_j with $e_j = |e_j| \exp(i\phi_j)$ and the model is $m(z) = \sigma(z)$, the Fréchet kernels are given by (Oldenburg 1979)

$$G_j(\sigma_0; z) = -\mu_0 \omega_j |e_j| \operatorname{Im}\left\{\frac{E(\sigma_0; z, \omega_j)^2}{E(\sigma_0; 0, \omega_j) E'(\sigma_0; 0, \omega_j)}\right\} \qquad (27)$$

for $|e_j|$, and

$$G_j(\sigma_0; z) = \mu_0 \omega_j \operatorname{Re}\left\{\frac{E(\sigma_0; z, \omega_j)^2}{E(\sigma_0; 0, \omega_j) E'(\sigma_0; 0, \omega_j)}\right\} \qquad (28)$$

for ϕ_j. If the model is $m(z) = \log \sigma(z)$, then $G_j(\sigma_0; z)$ in (27) and (28) is replaced by $\sigma_0(z) G_j(\sigma_0; z)$.

If the region $[0, a]$ is partitioned and a starting conductivity model $\sigma_0(z)$ specified, then an extremal conductivity model $\sigma(z)$ can be constructed by applying the LP formulation described in Sections 1 and 2 to the linearized equation (25). In practice, the inversion is iterated until the constructed extremal model adequately reproduces the observed data according to the χ^2 criterion (Parker 1977) and the extremal value of $\bar{\sigma}$ does not change significantly between iterations. If the procedure is initiated with an acceptable starting model, we have found the algorithm generally converges to a stable solution within two or three iterations. This likely reflects the fact that the inverse problem for conductance (depth-integrated conductivity) is well posed (Weidelt 1985). Moreover, we have found that the same solution is generally obtained when the algorithm is initiated with very different starting models; this provides some confidence that a global (rather than local) extremum for $\bar{\sigma}$ has been found.

To illustrate the appraisal in the analysis of MT data, we consider the following example. The true conductivity model consists of the test case considered by Whittall & Oldenburg (1989) in their survey of 1-D MT inversion techniques. The model, shown by the dashed line in Fig. 7(a and b), consists of four homogeneous layers overlying a uniform half-space. 26 data were generated at 13 periods equally spaced in logarithmic time from 0.0025 to 250 s.

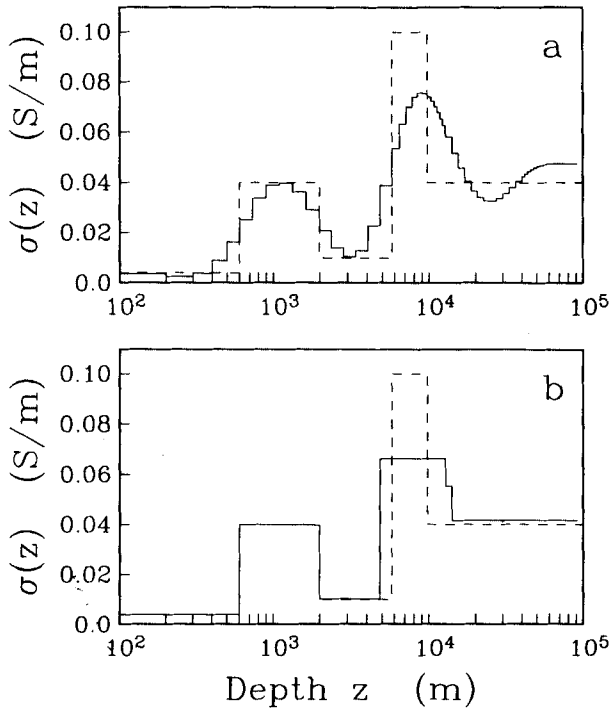

Figure 7. Minimum-structure conductivity models for the synthetic MT example. (a) shows the l_2 flattest model, (b) shows the l_1 minimum-variation model. The true model is indicated by the dashed line.

Although the data used in the inversions were accurate, they were assumed to have an uncertainty of 2.5 per cent in amplitude and phase so that a χ^2-misfit criterion could be used to measure the relative fit of the models.

Figure 7 shows two minimum-structure models constructed by utilizing (25) in an iterative algorithm. Fig. 7(a) shows the l_2 flattest model computed by minimizing the norm given by (26) with $m(z) = \sigma(z)$, $f(z) = z$ and $w(z) = 1$; Fig. 7(b) shows the minimum-variation model computed by minimizing the total-variation norm given by (21). Both constructed models reproduce the observed data to within a χ^2 misfit of 26. Minimizing the l_2 norm of the model gradient discriminates against large abrupt changes in conductivity and the model shown in Fig. 7(a) varies in a smooth, continuous manner with depth. Minimizing the l_1 norm of the total variation does not discriminate against abrupt changes, but produces a minimum-structure model that resembles a layered earth with structural variations occurring at distinct depths, as shown in Fig. 7(b).

The constructed models in Fig. 7 generally reproduce the features of the true model. In particular, both models indicate a region of low conductivity centred near 4000 m depth. We shall compute upper and lower bounds for conductivity averages $\bar{\sigma}(z_0 = 4000, \Delta)$ of this region. Fig. 8(a) shows the bounds computed when model limits $\sigma^- = 0.002$, $\sigma^+ = 0.2\ \mathrm{S\,m^{-1}}$ are imposed, but no constraint is placed on the total variation of the extremal models. The upper bound decreases from the imposed upper limit of $0.2\ \mathrm{S\,m^{-1}}$ for $\Delta \leq 300$ m to a value of about $0.047\ \mathrm{S\,m^{-1}}$ at $\Delta = 3000$ m. The computed lower bounds for regions of low conductivity are often not particularly meaningful since MT measurements contain little information about resistive

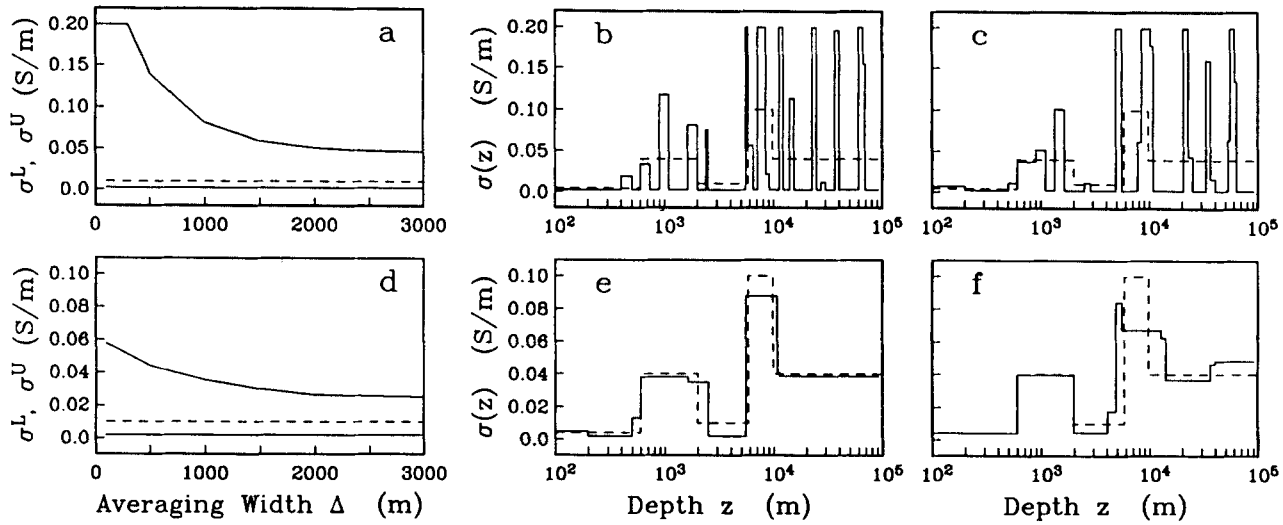

Figure 8. Lower and upper bounds for $\bar{\sigma}(z_0 = 4000, \Delta)$ are shown in (a) and (d). In (a) only model limits $0.002 \leq \sigma_i \leq 0.2\ \mathrm{S\,m^{-1}}$ have been imposed; in (d) a variation bound of $V_b = 0.21\ \mathrm{S\,m^{-1}}$ (the variation of the true model) was also included. The true model averages are indicated by the dashed line. The two plots to the right of each funnel function diagram show the constructed models which minimize and maximize $\bar{\sigma}(z_0 = 4000, \Delta = 3000)$. In these plots the true model is indicated by the dashed line.

layers. In this case the lower bound simply reflects the imposed lower limit over the entire width $100 \leq \Delta \leq 3000$ m.

Examples of extremal models which produce the funnel function bounds of Fig. 8(a) are given in Figs 8(b) and (c) which show, respectively, the constructed models which minimize and maximize $\bar{\sigma}(z_0, \Delta = 3000)$. These models have the characteristic sparse, spiky form of solutions of unconstrained variation, consisting of regions of conductivity at the imposed lower limit with isolated zones of high conductivity at or near the upper limit. All extremal models constructed to produce the bounds in Fig. 8(a) are of this form and have total variation values of 2.0 to $3.0\ \mathrm{S\,m^{-1}}$, more than 10 times that of the true model.

It is interesting to compare the LP extremal models with the theoretical results of Weidelt (1985). Weidelt treated the full non-linear problem of extremizing the conductance function $S(z_2) = \int_0^{z_2} \sigma(z)\, dz$ subject to exactly fitting a small number of MT data. He determined that when no model limits (except positivity) are imposed, the extremal models consist of insulating zones ($\sigma = 0$) and thin regions of infinite conductivity, but finite conductance, located at isolated points. When $S(z_2)$ is maximized, a conducting region is located at $z_2 - 0$, which is just included in the region of integration, whereas when $S(z_2)$ is minimized, a conducting region is located at $z_2 + 0$, which is just excluded. When model limits $\sigma^- \leq \sigma \leq \sigma^+$ are imposed, Weidelt found that the extremal models consist of a sequence of sections of alternating conductivities σ^- and σ^+. In general, when $S(z_2)$ is maximized, a layer of conductivity σ^+ ends at $z = z_2$, whereas when $S(z_2)$ is minimized, a layer of conductivity σ^+ begins at $z = z_2$. The extremal models of unconstrained variation in Fig. 8(b and c) (and in Figs 1 and 2 for the linear example) appear to be of generally the same form as Weidelt's exact solutions; the discrepancies are likely due to the finite discretization interval employed.

Weidelt's exact extremal models represent an interesting and important result. However, if the purpose of the extremization is to determine meaningful bounds for $\bar{\sigma}(z_0, \Delta)$, these types of models might not be satisfactory. Fig. 8(d) shows the result of imposing a variation bound of $V_b = 0.21\ \mathrm{S\,m^{-1}}$, which represents the total variation of the true model. A significant improvement in the resolution is apparent when Figs 8(a) and (d) are compared. The computed lower bound still reflects the imposed lower limit of $0.002\ \mathrm{S\,m^{-1}}$, however the upper bound has been reduced significantly, particularly at small widths. For $\Delta = 3000$ m bound for the average conductivity are $0.002 \leq \bar{\sigma} \leq 0.025\ \mathrm{S\,m^{-1}}$. The per cent improvement, $P(z_0, V_1, V_2, \Delta)$ given by (22), for $V_1 = \infty$ (no variation bound) and $V_2 = 0.21\ \mathrm{S\,m^{-1}}$ is shown in Fig. 9. For most averaging widths the funnel function bounds are improved by 50–75 per cent. Examples of the extremal models which produced the improved funnel function bounds are given in Fig. 8(e and f) which show the models which minimize and maximize $\bar{\sigma}(z_0, \Delta = 3000)$ for $V_b = 0.21\ \mathrm{S\,m^{-1}}$.

The constructed models shown in Fig. 7(a and b) exhibit a region of low conductivity centred at about 4000 m followed by a region of high conductivity centred at about 8000 m. To verify if this structure is required by all acceptable models we have computed upper bounds for $z_0 = 4000$ m and lower bounds for $z_0 = 8000$ m. Fig. 10(a) shows these bounds when

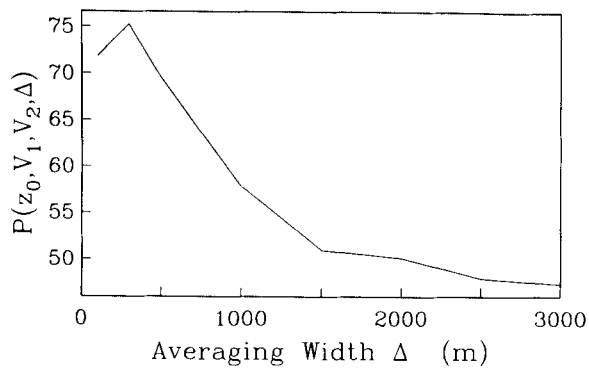

Figure 9. The percent improvement, P, for $V_1 = \infty$ (no variation bound) and $V_2 = 0.21\ S\ \mathrm{m^{-1}}$ for the synthetic MT example with $z_0 = 4000$ m.

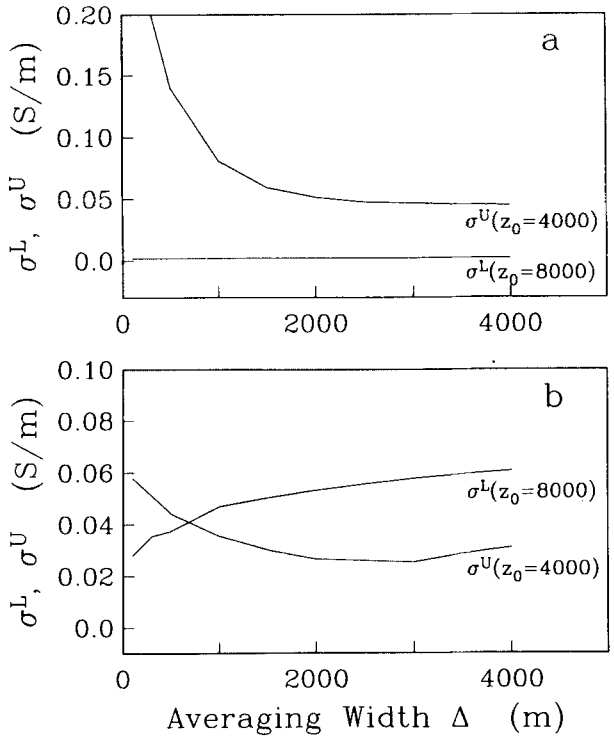

Figure 10. The upper bound $\sigma^U(z_0 = 4000, \Delta)$ is compared to the lower bound $\sigma^L(z_0 = 8000, \Delta)$ as a function of averaging width Δ. In (a) only model limits $0.002 \le \sigma_i \le 0.2\ \mathrm{S\ m^{-1}}$ were imposed, while in (b) a variation bound of $V_b = 0.21\ \mathrm{S\ m^{-1}}$ was also imposed.

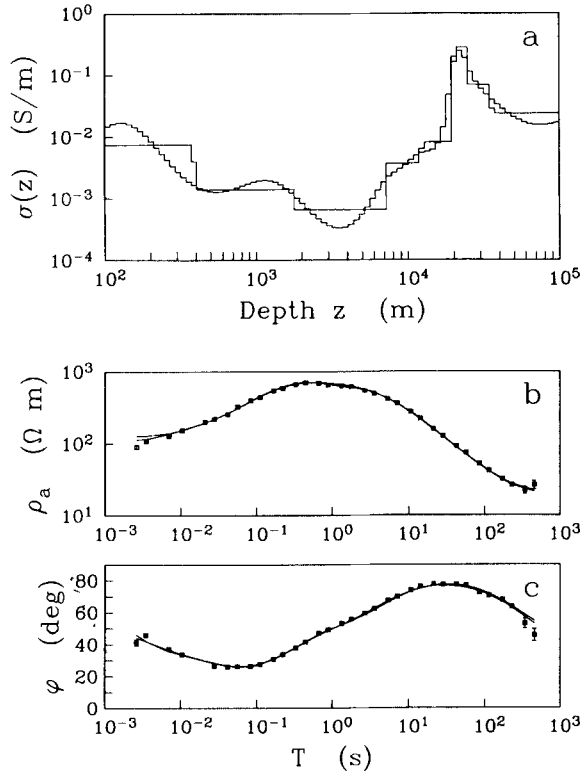

Figure 11. Minimum-structure models and MT responses observed in southeastern British Columbia, Canada. (a) The l_2 flattest model (smoothly varying curve) and the l_1 minimum-variation model; (b) and (c) comparison of the observed apparent resistivities and phases (squares) with those predicted from the constructed models (solid lines).

model limits $\sigma^- = 0.002$, $\sigma^+ = 0.2\ \mathrm{S\ m^{-1}}$ were imposed but the total variation was unconstrained. The computed lower bound for $z_0 = 8000$ m is lower than the upper bound for $z_0 = 4000$ m for all Δ, indicating that without additional information the difference between the two regions can not be resolved from the data. The results of including a variation bound $V_b = 0.21\ \mathrm{S\ m^{-1}}$ are shown in Fig. 10(b). For resolution widths greater than about 700 m the computed lower bound for $z_0 = 8000$ m is greater than the upper bound for $z_0 = 4000$ m indicating that the region of low conductivity followed by a region of higher conductivity is clearly resolved and is a required feature of all acceptable models with a total variation $V < V_b$.

6 MT FIELD DATA EXAMPLE

As a final illustration of our appraisal, we analyse a set of wide-band MT field data measured near Kootenay Lake in southeastern British Columbia, Canada. The data were collected as part of the LITHOPROBE Southern Cordilleran transect, and a preliminary analysis has been carried out by Jones *et al.* (1988). Fig. 11 shows the determinant-average MT responses (Ranganayaki 1984) together with two minimum-structure models. Jones *et al.* (1988) found that, with the exception of the single longest period, the data are consistent with the response of a one-dimensional model according to the criterion of Parker (1980) and Weidelt (1986). Fig. 11(a) shows the l_2 flattest model constructed by minimizing the gradient in log conductivity with log depth [i.e. $m(z) = \log \sigma(z)$, $f(z) = \log z$ and $w(z) = 1$ in (26)], and the l_1 minimum-variation model constructed by minimizing the total variation of $m(z) = \log \sigma(z)$. Figs 11(b) and (c) compare the observed apparent resistivities and phases with those predicted for the constructed models.

The two models shown in Fig. 11(a) are in good agreement and show essentially the same features as those obtained by Jones *et al.* (1988). In particular, the models indicate a region of low conductivity at 2000–7000 m depth and a region of high conductivity at 20 000–30 000 m depth. We will appraise these features.

In Fig. 11(a) we have considered our model to be $\log \sigma(z)$ rather than $\sigma(z)$ since we are interested in conductivity variations over several orders of magnitude. In order to appraise model features we still wish to determine bounds for the conductivity average $\bar{\sigma}$; however, to construct realistic extremal models of $\log \sigma(z)$, we need to constrain the total variation of $\log \sigma(z)$. This can be accomplished as follows.

If the model $m(z)$ is taken to be $\log \sigma(z)$, the definition of the total variation given by (10) becomes

$$V[\log \sigma] = k \int_0^a |d\sigma/\sigma|,$$

where $k = \log e$. In discrete form, this can be approximated by

$$V[\log \sigma] = k \sum_{i=1}^{n-1} \frac{|\Delta\sigma_i|}{(\sigma_{i+1} + \sigma_i)/2}. \qquad (29)$$

The variation of $\log \sigma(z)$ can be constrained by applying the LP formulation of Method 2, Section 2. The variational constraint (20) is modified to be

$$k \sum_{i=1}^{n-1} \frac{r_i + s_i}{(\sigma_{i+1} + \sigma_i)/2} \leq V_b \tag{30}$$

where σ_i, σ_{i+1} are values of the conductivity model of the previous iteration and V_b is a bound on the log variation. Since (29) is not an exact representation of the variation of $\log \sigma(z)$, the actual log variation of the constructed model, $\Sigma_i(\log \sigma_{i+1} - \log \sigma_i)$, must be evaluated directly.

Since the log variation constraint depends on the model of the previous iteration, it is important that the iterations converge to a stable solution where changes in the model between iterations are negligible. To ensure that differences between successive iterations are not too great, we have limited changes in the model to a factor of 2. Although this may not be an optimal procedure for stabilizing the convergence, it has proved successful for all cases we have considered. We have also found that the same solution is obtained when the algorithm is initiated with different starting models; this provides confidence that the extremal value of $\bar{\sigma}$ is independent of $\sigma_0(z)$.

Figure 12(a) shows the extremal model which maximizes $\bar{\sigma}$ over the apparent low conductivity region, 2000–7000 m depth. Conductivity limits $\sigma^- = 0.0001$, $\sigma^+ = 1.0\ \mathrm{S\,m^{-1}}$ were imposed, but no bound was included on the variation. The computed upper bound for $\bar{\sigma}$ is $0.0024\ \mathrm{S\,m^{-1}}$ and the log variation of the model is 76. The model is sparse and spiky, consisting of regions of conductivity at the imposed lower limit with isolated zones of high conductivity. A narrow zone of high conductivity (one partition element wide) is just included at each edge of the region of maximization. Such a model is not appealing from a geophysical point of view.

Figures 12(b) and (c) show extremal models with log variations of 15 and 9.2, respectively. The rapid fluctuations between low and high conductivity values have been suppressed. The upper bounds for $\bar{\sigma}$ computed from the models in Figs 12(b) and (c) are 0.0020 and $0.0017\ \mathrm{S\,m^{-1}}$. The models which minimize $\bar{\sigma}$ for this region simply reflect the imposed lower limit and are not shown.

In many practical cases, an appropriate *a priori* bound for the variation may not be known. When this is the case, the interpreter may wish to construct extremal models for a number of variation bound values and select the model with

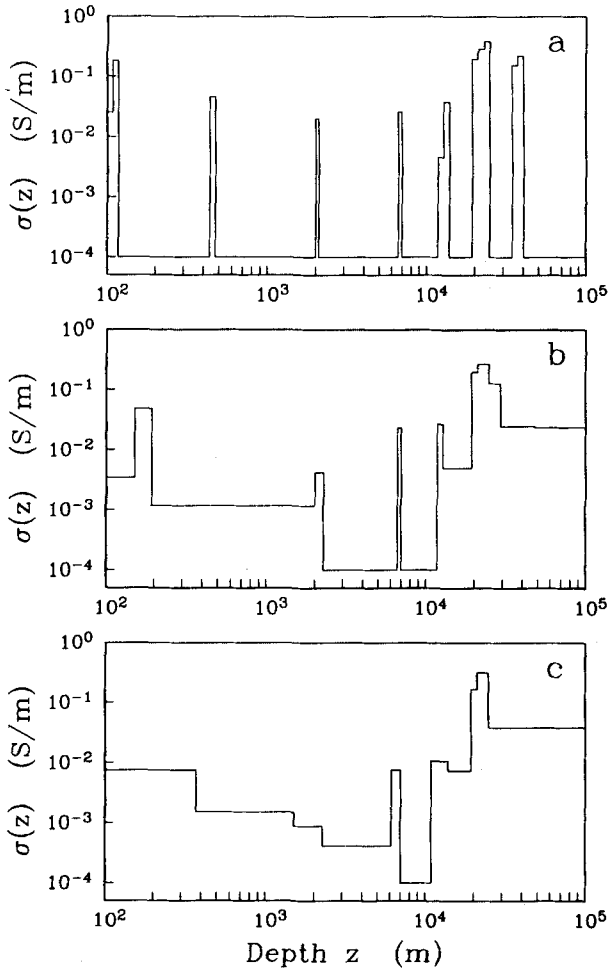

Figure 12. Constructed models which maximize $\bar{\sigma}$ for the apparent low conductivity region, 2000–7000 m depth. Model limits $\sigma^- = 0.0001$, $\sigma^+ = 1.0\ \mathrm{S\,m^{-1}}$ were imposed in each case. (a) shows the extremal model of unconstrained variation; (b) and (c) show extremal models with log variations of 15 and 9.2, respectively.

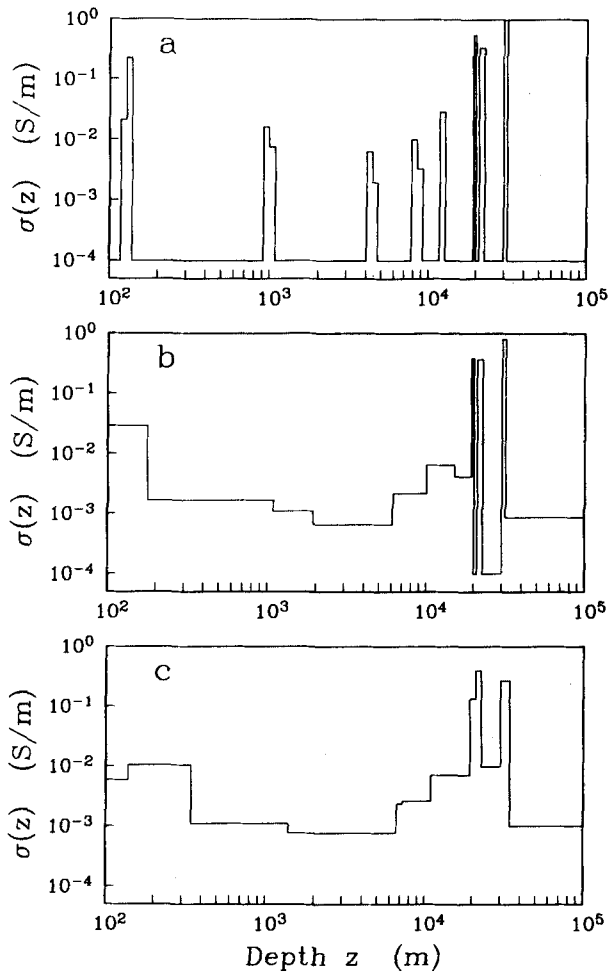

Figure 13. Constructed models which minimize $\bar{\sigma}$ for the apparent high conductivity region, 20 000–30 000 m depth. Model limits $\sigma^- = 0.0001$, $\sigma^+ = 1.0\ \mathrm{S\,m^{-1}}$ were imposed in each case. (a) shows the extremal model of unconstrained variation; (b) and (c) show extremal models with log variations of 24 and 9.5, respectively.

the largest variation that is deemed geophysically plausible. For example, the extremal model shown in Fig. 12(a) is not realistic; however, the model shown in Fig. 12(c) might be considered acceptable and therefore a meaningful bound for $\bar{\sigma}$ would be $0.0017\,\mathrm{S\,m^{-1}}$. In this particular example where the computed upper bound does not change greatly with variation, constructing extremal models with a number of different variations also serves to verify that a meaningful bound has been determined.

Extremal models which minimize and maximize $\bar{\sigma}$ for the apparent high conductivity region at 20 000–30 000 m depth are shown in Figs 13 and 14. Model limits $\sigma^- = 0.0001$, $\sigma^+ = 1.0\,\mathrm{S\,m^{-1}}$ were imposed in each case. Fig. 13(a) shows the model of unconstrained variation which minimizes $\bar{\sigma}$. The lower bound for $\bar{\sigma}$ computed from this model is $0.060\,\mathrm{S\,m^{-1}}$ and the log variation is 78. Fig. 13(b and c) shows minimization models with log variations of 24 and 9.5. The lower bounds for $\bar{\sigma}$ computed from these models are 0.066 and $0.091\,\mathrm{S\,m^{-1}}$, respectively. The extremal model of unconstrained variation which maximizes $\bar{\sigma}$ is shown in Fig. 14(a). The upper bound for $\bar{\sigma}$ computed from this model is $0.30\,\mathrm{S\,m^{-1}}$ and the log variation is 75. Fig. 14(b and c) shows maximization models with log variations of 24 and 8.9; the computed upper bounds for $\bar{\sigma}$ are 0.30 and $0.26\,\mathrm{S\,m^{-1}}$, respectively.

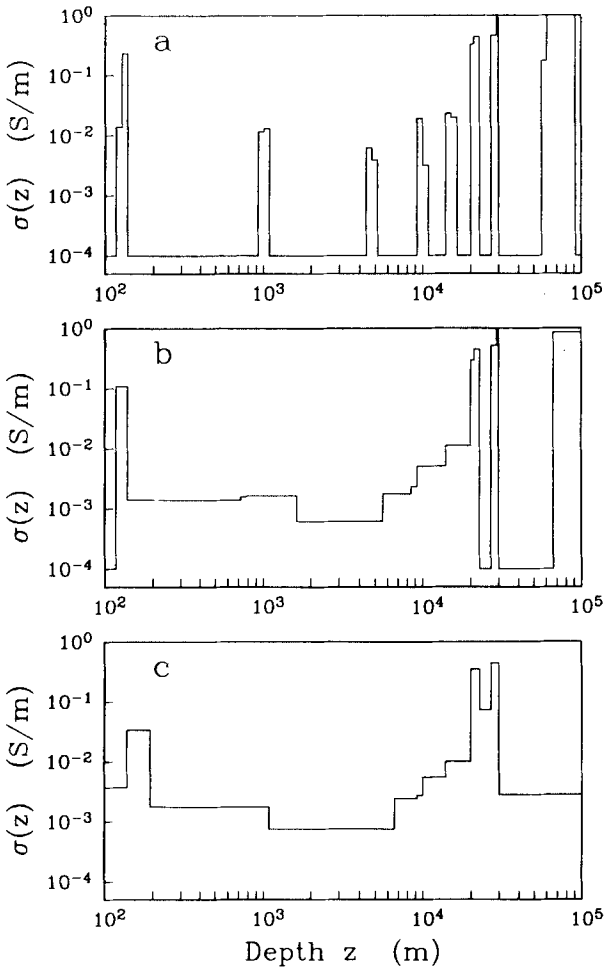

Figure 14. Constructed models which maximize $\bar{\sigma}$ for the apparent high conductivity region, 20 000–30 000 m depth. Model limits $\sigma^- = 0.0001$, $\sigma^+ = 1.0\,\mathrm{S\,m^{-1}}$ were imposed in each case. (a) shows the extremal model of unconstrained variation; (b) and (c) show extremal models with log variations of 24 and 8.9, respectively.

If we accept the extremal models shown in Figs 13(c) and 14(c) as geophysically realistic representations of the Earth, bounds for the average conductivity are $0.091 \leq \bar{\sigma} \leq 0.26\,\mathrm{S\,m^{-1}}$. This establishes the region at 20 000–30 000 m depth as a zone of high conductivity. The average conductivity of this is greater than that of the low conductivity zone at 2000–7000 m depth ($\bar{\sigma} \leq 0.0017\,\mathrm{S\,m^{-1}}$) by at least an order of magnitude.

7 DISCUSSION

In this paper we have shown that model features can be appraised by constructing extremal models to determine bounds for averages of the model over specified regions. In order to compute optimal bounds, it is important that the extremal models be geophysically realistic. Unfortunately, extremizing the model average while still satisfying the data constraints often produces models with an excessive amount of structure. In many practical applications we are not willing to accept such models even if they are consistent with the observed data. In order to incorporate this prejudice, we have introduced the total variation as a measure of structure and have presented two methods of bounding the variation to produce reasonable extremal models. Restricting the variation discriminates against highly oscillatory models and, as consequence of operating on a restricted model space, the difference between the upper and lower bounds is often considerably reduced.

It would be advantageous if the variation bound V_b could be ascertained through analysis or from the physics of the problem. Unfortunately, this is seldom the possible. Nevertheless, there are several ways to proceed. One way is to specify V_b as some multiplicative factor of the variation of the least-structure model. Alternatively, in a more complete analysis we construct models for a number of values of V_b ranging from the minimum to that of the unbound extremal model, and select the model with the largest variation that is judged to be geophysically plausible. This establishes V_b for the analysis. In this manner the interpreter makes use of any knowledge or insight regarding the variation of the model to select the reasonable extremal models and meaningful funnel function bounds. Lastly, even though the constructed models vary greatly with V_b, it is often found that approximately the same bounds for $\bar{m}$ are computed for a wide range of V_b. This provides confidence that even when a non-optimal variation bound is used, meaningful bounds for $\bar{m}$ may still be computed.

The ability to estimate bounds for averages of the model depends crucially upon the restructuring of the linearized data equations (see equation 25). This linearized approach, however, raises concern about whether the algorithm converges to a global extremum. We know of no way to evaluate this directly, and it is not clear whether this transformation will prove equally successful in all non-linear problems. For the MT application we have found that the same solution is obtained when the algorithm is initiated with very different starting models. This does not constitute proof, but it does provide some confidence that a global extremum has been found.

In the MT application, because the conductivity σ can

vary over many orders of magnitude, it is often desirable to evaluate the geologic reasonableness of extremal models in a log–log domain. This adds a further complication. We have introduced a method of bounding $V[\log \sigma]$ which requires using the conductivity model of the previous iteration. Although this slows the convergence somewhat, by initiating the algorithm with an acceptable minimum-structure model and limiting the changes in the model between successive iterations, we have found the algorithm to work effectively.

ACKNOWLEDGMENTS

This research was funded by NSERC grant 5-84270. S.E.D. was supported by an NSERC Postgraduate Scholarship.

REFERENCES

Backus, G. E., 1970a. Inference from inadequate and inaccurate data, 1, *Proc. natn. Acad. Sci. USA*, **65,** 1–7.

Backus, G. E., 1970b. Inference from inadequate and inaccurate data, 2, *Proc. natn. Acad. Sci. USA*, **65,** 281–287.

Backus, G. E., 1970c. Inference from inadequate and inaccurate data, 3, *Proc. natn. Acad. Sci. USA*, **67,** 282–289.

Backus, G. E., 1972. Inference from inadequate and inaccurate data, in *Mathematical Problems in Geophysical Sciences*, American Mathematical Society, Providence, RI.

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

Constable, S. C., Parker, R. L. & Constable, C. G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data, *Geophysics*, **52,** 289–300.

Heustis, S. P., 1987. Construction of non-negative resolving kernels in Backus–Gilbert theory, *Geophy. J. R. astr. Soc.*, **90,** 495–500.

Jones, A. G. *et al.*, 1988. Magnetotelluric observations along the LITHOPROBE Canadian Cordilleran Transect, *Geophys. Res. Lett.*, **15,** 667–680.

Korevaar, J., 1968. *Mathematical Methods*, Vol. 1, Academic Press, New York.

Lang, S. W., 1985. Bounds from noisy linear measurements, *IEEE Trans. Inform. Theory*, **IT-31,** 490–508.

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., 1984. An introduction to linear inverse theory, *IEEE Trans. Geosci. Remote Sensing*, **GE-22,** 665–674.

Parker, R. L., 1974. Best bounds on density and depth from gravity data, *Geophysics*, **39,** 644–649.

Parker, R. L., 1975. The theory of ideal bodies for gravity data interpretations, *Geophys. J. R. astr. Soc.*, **42,** 315–334.

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

Parker, R. L., 1980. The inverse problem of electromagnetic induction: existence and construction of solutions based on incomplete data, *J. geophys. Res.*, **85,** 4421–4425.

Ranganayaki, R. P., 1984. An interpretive analysis of magnetotelluric data, *Geophysics*, **49,** 1730–1748.

Smith, J. T. & Booker, J. R., 1988. Magnetotelluric inversion for minimum structure, *Geophysics*, **52,** 1565–1576.

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

Weidelt, P., 1986. Discrete frequency inequalities for magnetotelluric impedances of one-dimensional conductors, *J. Geophys.*, **59,** 171–176.

Whittall, K. P. & Oldenburg, D. W., 1990. Inversion of magnetotelluric data over a one-dimensional earth, in *Magnetotellurics in Geophysical Exploration*, ed. Wannamaker, P. E., SEG publication.