

# Inversion of geophysical data using an approximate inverse mapping

D. W. Oldenburg and R. G. Ellis

*Department of Geophysics and Astronomy, University of British Columbia, Vancouver, Canada V6T 1W5*

Accepted 1990 November 6. Received 1990 November 5; in original form 1990 April 20

## SUMMARY

Realistic geologic features are 3-D and inverse techniques which rely upon linearization and computation of a sensitivity matrix to show how a change in the model affects a particular datum, can require prohibitive amounts of computation. Even  $10^4$  data collected over an earth parametrized into  $10^2 \times 10^2 \times 10^2$  elements has a sensitivity matrix which is  $10^4 \times 10^6$ . The generation of that matrix requires the solution of many 3-D forward problems and its solution is also computationally intensive. In this paper we formulate a general technique for solving large-scale inverse problems which does not involve full linearization and which can obviate the need to solve a large system of equations. The method uses accurate forward modelling to compute responses, but only uses an approximate inverse mapping to map data back to model space. The approximate inverse mapping is chosen with emphasis on the physics of the problem and on computational expediency. There are two ways to implement the AIM (Approximate Inverse Mapping) inversion. At any iteration step, AIM-MS applies the approximate inverse mapping to forward modelled data and also applies the same mapping to the observations; the model perturbation is taken as the difference between the resulting functions. In AIM-DS, an alteration to the data is sought, such that the approximate inverse mapping applied to the altered data yields a model which adequately satisfies the observations. The approximate mapping inversion is illustrated with a simple parametric inverse problem and with the inversion of magnetotelluric (MT) data to recover a 1-D conductivity model. To illustrate the technique in a realistically complicated problem we invert MT data acquired from a line of stations over a 2-D conductivity structure. TE and TM mode data are inverted individually and as determinant averages. As a final example we invert 900 data, with and without noise, to recover a model that is parametrized by 1500 cells of unknown conductivity. The inversion is found to be computationally efficient and robust.

**Key words:** inverse problems, magnetotellurics.

## 1 INTRODUCTION

Geophysical inverse problems are usually non-linear, ill-posed and often large-scale. The non-linearity is inherent in the mathematical relations used to describe the physical experiment, while the ill-posedness is caused by Irrelevant and Unimportant parameters (Jupp & Vozoff 1975), or sometimes by invalid assumptions used in defining the space of possible models used in the solution. The large-scale nature is the result of the real world applications which geophysical inverse problems attempt to address. Consequently, the solution of the geophysical inverse problem is non-trivial and has been attacked by several methods with varying degrees of success.

There are several standard methods used in the solution of non-linear inverse problems with perhaps the two best known being the gradient method and the quasi-Newton method. When the inverse problem is also ill-posed a regularization scheme is required (Tikhonov & Arsenin 1977); the more common regularization techniques use a spectral expansion of the sensitivity matrix and eliminate Unimportant and Irrelevant parameters by truncation, or reduce their effects using a damped least-squares method (Levenberg 1944; Marquardt 1970). Successful derivatives of the damped least-squares technique involve modification of the objective function to be minimized. These methods of solution for the non-linear ill-posed inverse problem can provide excellent solutions to many smaller scale problems;

however, they are computationally intensive and only marginally applicable to large-scale inverse problems.

The large-scale nature of geophysical inverse problems and the high computational cost of existing methods encouraged us to seek an alternative inversion scheme. Our basic philosophy was that knowledge of the physics of the problem must be introduced in an optimal manner and wherever possible used to constrain the inverse problem and consequently reduce the computational cost. This differs from the standard methods which typically use mathematical techniques (e.g. Taylor series expansion) to simplify the problem. Instead we try to break the physics into several levels of importance with a division between the primary physical process and the residual physical effects. For example, in the 2-D or 3-D magnetotelluric problem the primary physical process is usually the 1-D attenuation of the magnetic and electric fields as they propagate through the conductive earth. The residual physical effects may be the scattering of the fields caused by 2-D or 3-D conductivity distributions.

The philosophy of primary and residual physics leads naturally to the concepts of exact and approximate mappings. An exact forward mapping,  $F$ , by definition, maps from a space of models,  $H_m$ , into a space of data,  $H_e$ , using the complete physics of the system. Referring back to the magnetotelluric example, the exact mapping is derived from Maxwell's equations and maps a conductivity distribution into an impedance tensor. An approximate forward mapping,  $\tilde{F}$ , maps model space into data space using only part of the physics of the system. Again, for the 2-D or 3-D magnetotelluric example, where impedances are to be computed at a number of sites on the earth surface, then a 1-D computation for the impedance at each site using the conductivity vertically under that site, would be an approximate forward mapping.

In a similar manner we introduce exact and approximate inverse mappings. Ideally, one desires an exact inverse mapping,  $F^{-1}$  such that

$$FF^{-1}(e^{\text{obs}}) = e^{\text{obs}} \quad (0)$$

where  $e^{\text{obs}}$  are the observed data and  $\approx$  allows for discrepancies due to noise in the data. Unfortunately,  $F^{-1}$  is not usually available. This leads us to introduce an approximate inverse mapping,  $\tilde{F}^{-1}$ , which produces a model  $\tilde{F}^{-1}(e^{\text{obs}})$ . Generally  $\tilde{F}^{-1}$  is constructed in accordance with the primary physics of the problem. For instance, in the 2-D magnetotelluric problem  $\tilde{F}^{-1}$  might be defined as a series of 1-D inversions. Application of the forward mapping to the constructed model yields predicted data

$$F\tilde{F}^{-1}(e^{\text{obs}}) = e^p. \quad (1)$$

Most likely the difference between  $e^{\text{obs}}$  and  $e^p$  is sufficiently large that the constructed model would not be considered acceptable. With proper iterative improvement however, it may be possible to use  $\tilde{F}^{-1}$  to converge to a model which reproduces the data to within some acceptable error tolerance. In this paper we develop two algorithms to accomplish this. At the heart of each of these algorithms is an Approximate Inverse Mapping and we refer to our procedures as AIM inversions. The algorithms differ in the space in which modification is carried out. In the first

algorithm, an element in  $H_m$  is updated and we refer to this as AIM-MS where MS is an acronym for Model Space. In the second algorithm, an element in  $H_e$  is updated; we refer to this as AIM-DS where DS indicates Data Space.

The concept of developing iterative inversion algorithms by exploiting the properties of the approximate inverse mapping has been presented before. In their 1-D inversion of MT data Goldberg, Loewenthal, and Rotstein (1982) computed a perturbation in model space by differencing the models obtained after applying the approximate inverse mapping to observed and forward modelled data; their algorithm is an example of an AIM-MS inversion. Zohdy (1974) inverts dc resistivity data using Dar Zarrouk curves and successively updates the data; his algorithm is of the AIM-DS type. No doubt there are other examples in the literature where these ideas have been used. The importance of the work in this paper however is to formalize the use of approximate inverse mappings and to show how they can be derived and used successfully to solve large-scale geophysical inverse problems.

This paper is divided into four major sections. We first derive the basic AIM algorithms. Particular emphasis is placed on the duality of perturbations in  $H_m$  and  $H_e$  and upon quantification of a mapping error defined in each of those spaces. In the next section we illustrate the AIM algorithms with a simple parametric example of relevance to electromagnetic induction. This example sets a foundation for understanding the essential components of the AIM inversions even though model space and data space are both 2-D. The next section illustrates how the AIM technique can be used to recover a 1-D conductivity depth function which accurately reproduces surface measured magnetotelluric responses. In the final section we present examples of 2-D magnetotelluric inversions which may ultimately serve as a *template for solving large-scale 3-D geophysical problems*. There we use an approximate inverse mapping based upon 1-D inversions to recover a 2-D electrical conductivity structure that reproduces magnetotelluric data acquired from a line of stations. The geologic models consist of one and two prisms buried in a two-layer earth. Both TE and TM mode data are inverted individually, and after being combined into a determinant average (Berdichevsky & Dmitriev 1976).

## 2 APPROXIMATE MAPPING INVERSION

There are at least two ways to derive the recursive solution for the AIM-MS algorithm. Although the final formulae will be shown to be identical, the justification for presenting both is that they provide different insight into the inversion algorithm. In the first derivation, the model is built up through successive additions of perturbations whose magnitude depends upon the difference between the observations and the forward modelled data at the current iteration. This is perhaps the viewpoint that is most closely related to linearized methods. The second derivation is formulated directly from the definition of a mapping error defined on model space.

Let  $m^{(0)}$  denote the starting model and let  $m^{(n)}$  be the model constructed at the  $n$ th iteration. Let the forward and approximate inverse mappings be denoted by  $F$  and  $\tilde{F}^{-1}$  respectively. Application of the forward mapping to  $m^{(n)}$

produces the predicted responses

$$F[m^{(n)}] = e^{(n)}. \quad (2)$$

Application of an approximate inverse mapping  $\tilde{F}^{-1}$  yields a model denoted by a tilde:

$$\tilde{F}^{-1}[e^{(n)}] = \tilde{m}^{(n)}. \quad (3)$$

When the approximate inverse mapping is applied to the observations, the result is  $\tilde{m}^{\text{obs}}$ .

Because the inverse mapping is approximate, it is not expected that application of  $F$  to any model  $m \in H_m$  followed by application of  $\tilde{F}^{-1}$  will reproduce the initial model. The discrepancy can be quantified in terms of a model space mapping error defined as

$$\Delta_m(m) \equiv m - \tilde{F}^{-1}F(m) = (\mathcal{J}_m - \tilde{F}^{-1}F)(m), \quad (4)$$

where  $\mathcal{J}_m$  is the identity mapping on  $H_m$ . By using the definitions in (2) and (3), the mapping error for the  $n$ th model can be written as

$$\Delta_m[m^{(n)}] = m^{(n)} - \tilde{m}^{(n)}. \quad (5)$$

All of the above mappings, data, and models are illustrated in Fig. 1. This figure is fundamental to the development of the AIM-MS procedure and it provides considerable insight into understanding the AIM approach.

To formulate the inversion steps we proceed as follows. At the  $n$ th iteration, the predicted responses  $e^{(n)}$  are not in sufficient agreement with the data, and a model perturbation  $\delta m^{(n)}$  is desired which, when added to  $m^{(n)}$ , reproduces the data. That is

$$m^{(n+1)} = m^{(n)} + \delta m^{(n)} \quad (6)$$

and

$$F[m^{(n+1)}] = e^{\text{obs}}. \quad (7)$$

The mapping error  $\Delta_m[m^{(n+1)}]$ , obtained by combining equations (4) and (7), is

$$\Delta_m[m^{(n+1)}] = m^{(n+1)} - \tilde{m}^{\text{obs}}. \quad (8)$$

Applying equation (6) yields

$$\delta m^{(n)} = \tilde{m}^{\text{obs}} - m^{(n)} + \Delta_m[m^{(n+1)}]. \quad (9)$$

Unfortunately  $\Delta_m[m^{(n+1)}]$  is not known, but if we assume

that the mapping errors at the  $n$ th and  $(n+1)$ th models are approximately the same, that is, if  $\Delta_m[m^{(n+1)}] \approx \Delta_m[m^{(n)}]$ , then the final perturbation, obtained by using (5) to evaluate (9), is given by

$$\delta m^{(n)} = \tilde{m}^{\text{obs}} - \tilde{m}^{(n)} \quad (10)$$

and hence

$$m^{(n+1)} = m^{(n)} + [\tilde{m}^{\text{obs}} - \tilde{m}^{(n)}]. \quad (11)$$

The appropriate model perturbation is therefore the difference in model space between the application of  $\tilde{F}^{-1}$  to the observations and to the data predicted by the current model.

It may have been possible to intuit this result directly from Fig. 1. That figure suggests that under the application of the approximate inverse mapping, there is a 1-1 relationship between a specific change in data space and a specific change in model space. The misfit in data space,  $e^{\text{obs}} - e^{(n)}$ , is mapped to a model difference,  $\tilde{m}^{\text{obs}} - \tilde{m}^{(n)}$ . Using these discrete differences, a new model which adjusts for the data misfit discrepancy, is given by (11). It is perhaps this insight that has prompted previous use of equation (11). For example, Goldberg, Loewenthal & Rotstein (1982) use this model space update when inverting MT data. The advantage of our formal derivation lies in the explicit introduction of the approximate inverse mapping and the associated model space mapping error, and the development of a necessary condition for convergence of the iterative algorithm. That convergence condition requires that the difference between the mapping errors at successive iterations must tend to zero.

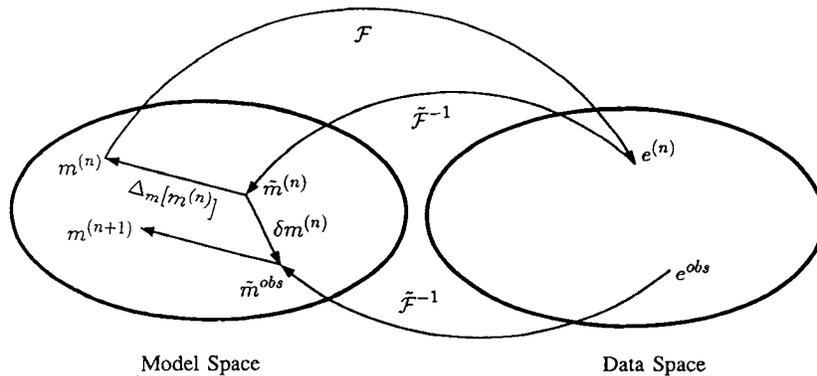
The second derivation for the AIM-MS algorithm can also be observed from Fig. 1. By definition of the mapping error given in equation (4), any model in  $H_m$  can be written as

$$m = \tilde{m} + \Delta_m(m). \quad (12)$$

If  $m$  is a solution to the inverse problem, that is, if  $F(m) = e^{\text{obs}}$ , then (12) can be written as

$$m = \tilde{m}^{\text{obs}} + \Delta_m(m). \quad (13)$$

However, since  $m$  is not known, iteration is used. Substituting for the current (known) model on the



**Figure 1.** A diagrammatic representation of the AIM-MS inversion. The inversion starts with  $\tilde{F}^{-1}$  applied to  $e^{\text{obs}}$  to produce a first model estimate  $\tilde{m}^{\text{obs}}$ . Either the mapping error  $\Delta_m[m^{(n)}]$  at the current model,  $m^{(n)}$ , or the perturbation  $\delta m^{(n)}$  are used to produce the  $(n+1)$ th model.

right-hand side yields an updated model

$$m^{(n+1)} = \bar{m}^{obs} + \Delta_m[m^{(n)}]. \tag{14}$$

Equation (14) is the desired iterative solution to the inverse problem. It has great appeal, not only because of its computational simplicity, but also because of the physical insight that it conveys. It shows that the final solution is composed of two parts. The first is  $\bar{m}^{obs}$ . Loosely speaking, we think of the exact mapping  $F$  as containing the primary and residual physics of the problem but think of  $\bar{F}^{-1}$  as containing only the primary physics. As a result  $\bar{m}^{obs}$  can be thought of as that part of the model which is consistent with the primary physics. The second part of equation (14) is a remainder needed to account for the neglect of the residual physics in the inverse mapping. We note that  $\Delta_m = (\mathcal{J}_m - \bar{F}^{-1}F)$  has the form of a projection operator. Its application to any element of  $H_m$  quantifies that part of the model which is annihilated by the approximate inverse mapping. The two terms in (14) therefore bring to the model both primary and residual physics contributions even though  $\bar{F}^{-1}$  is concerned only with the former.

The two recursive solutions in (11) and (14) are, of course, identical. Writing out the mapping error in (14) yields

$$m^{(n+1)} = \bar{m}^{obs} + [m^{(n)} - \bar{m}^{(n)}]. \tag{15}$$

This expression is identical to equation (11) except for the way in which the elements are grouped. The two ways of thinking about the AIM-MS solution: (i) generating a sequence of model perturbations based upon a specialized finite difference, and (ii) generating a model made up of a portion which is sensitive to the primary physics and adding a corrective term due to the model space mapping error, thus lead to the same solution.

The question of convergence aside, there is an important point to emphasize. The AIM-MS inversion is iterative but it does not necessarily involve linearization. This means there is no need to generate a sensitivity matrix or a Fréchet derivative to relate a small change in the model to a small change in the data. There is also no requirement in AIM-MS that differences in the data or differences in the model be small. When a model perturbation is constructed using equation (10) there is no necessity at the early

iterations that the difference between  $e^{obs}$  and  $e^{(n)}$  be small, even though, as convergence is achieved,  $e^{(n)}$  tends to  $e^{obs}$  and consequently  $\delta m^{(n)} \rightarrow 0$ . Also, in (14) there is no requirement that the final mapping error be small, in fact  $\|\Delta_m[m^{(n)}]\|$  will often be of comparable size to  $\|\bar{m}^{obs}\|$ . What is required in both derivations of AIM-MS is that  $\|\Delta_m[m^{(n+1)}] - \Delta_m[m^{(n)}]\| \rightarrow 0$ . This requires that small changes in the data, map to small changes in the model under the approximate inverse mapping; that is, the approximate inverse mapping must be stable.

In AIM-MS the final model is generated by successive additions of perturbation functions. Without any smoothing or regularization, we generally find that the model continues to increase in complexity and we have no explicit control over the type of model that is generated. It is not possible within the AIM-MS formalism to construct directly a model which minimizes a global objective function of the model and this has motivated a different formulation of the AIM inversion. The goal in AIM-DS is to find a ‘correction’ to the observed data so that when  $\bar{F}^{-1}$  is applied to the corrected data, the resultant model reproduces the observed data when it is operated upon by the exact forward mapping. That is, the inverse problem is solved by finding a data perturbation such that a new datum  $\bar{e} = e^{obs} + \phi$  satisfies

$$F\bar{F}^{-1}(\bar{e}) = F\bar{F}^{-1}(e^{obs} + \phi) = e^{obs}. \tag{16}$$

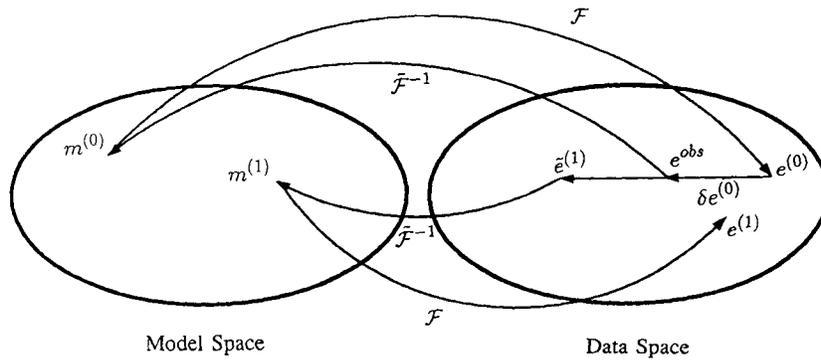
The advantage of this approach is that the desired characteristics of the final model may be built directly into  $\bar{F}^{-1}$  since the final model is computed directly by a single application of  $\bar{F}^{-1}$ . Again there are two ways to derive the recursive solution. We present both because different insight is involved.

The first method generates  $\phi$  as a sequence of additive perturbations. The basic procedure is illustrated in Fig. 2. Application of the approximate inverse mapping to the observations yields

$$m^{(0)} = \bar{F}^{-1}(e^{obs}). \tag{17}$$

An exact forward mapping applied to  $m^{(0)}$  yields the predicted data

$$e^{(0)} = F[m^{(0)}]. \tag{18}$$



**Figure 2.** A diagrammatic representation of the AIM-DS inversion. The inversion starts with  $\bar{F}^{-1}$  applied to  $e^{obs}$  to produce a first model estimate  $m^{(0)}$ . An exact forward mapping is applied to yield predicted data  $e^{(0)}$ . The difference between  $e^{obs}$  and  $e^{(0)}$  is an estimate of the data space mapping error. The predicted data are corrected by this difference to yield a new datum  $\bar{e}^{(1)}$ . The updating of data is continued recursively.

The first estimate for the correction to the data is

$$\phi^{(0)} = \delta e^{(0)} = e^{\text{obs}} - e^{(0)} \quad (19)$$

and the adjusted data are

$$\bar{e}^{(1)} = e^{\text{obs}} + \phi^{(0)}. \quad (20)$$

Application of  $F\bar{F}^{-1}$  to  $\bar{e}^{(1)}$  yields the predicted data  $e^{(1)}$  which are likely closer to  $e^{\text{obs}}$  than were  $e^{(0)}$ , but the discrepancy is probably large enough that a further adjustment is required. The total perturbation is built up cumulatively and the recursive equations are

$$\begin{aligned} m^{(n+1)} &= \bar{F}^{-1}[\bar{e}^{(n+1)}], \\ \bar{e}^{(n+1)} &= e^{\text{obs}} + \phi^{(n)}, \\ \phi^{(n)} &= \phi^{(n-1)} + \{e^{\text{obs}} - F\bar{F}^{-1}[\bar{e}^{(n)}]\}. \end{aligned} \quad (21)$$

Combining the last two equations yields

$$\begin{aligned} \bar{e}^{(n+1)} &= \bar{e}^{(n)} + e^{\text{obs}} - F\bar{F}^{-1}[\bar{e}^{(n)}] \\ &= \bar{e}^{(n)} + [e^{\text{obs}} - e^{(n)}]. \end{aligned} \quad (22)$$

The above derivation is analogous to the first derivation for the AIM-MS. Here however, it is the final data which are built up from a sequence of additive perturbations. We might expect therefore that these perturbed data will eventually exhibit enhanced structure as the iterations proceed.

In AIM-MS, we introduced a companion derivation based upon the mapping error in model space. We carry out an analogous derivation for AIM-DS. Define a mapping error on data space as

$$\Delta_e(e) \equiv (\mathcal{J}_e - F\bar{F}^{-1})(e) \quad (23)$$

where  $\mathcal{J}_e$  is the identity map. Applying equation (23) to  $\bar{e}^{(n)}$  yields

$$\bar{e}^{(n)} = e^{(n)} + \Delta_e[\bar{e}^{(n)}]. \quad (24)$$

Since our goal is to find an  $\bar{e}^{(n)}$  such that  $e^{(n)} = e^{\text{obs}}$ , the desired form of equation (24) is

$$\bar{e}^{(n)} = e^{\text{obs}} + \Delta_e[\bar{e}^{(n)}] \quad (25)$$

Because the left- and right-hand sides of equation (25) contain the desired function, we alter it to an iterative equation

$$\bar{e}^{(n+1)} = e^{\text{obs}} + \Delta_e[\bar{e}^{(n)}] \quad (26)$$

which is valid so long as  $\Delta_e[\bar{e}^{(n+1)}] \approx \Delta_e[\bar{e}^{(n)}]$  i.e.  $\bar{e}^{(n)}$  converges to a fixed point of  $e^{\text{obs}} + \Delta_e$ . Equation (26) represents the final form for the recursive relation. It shows explicitly that the desired perturbation  $\phi$  in equation (16) is the data space mapping error. Using the definition of  $\Delta_e$  we obtain

$$\bar{e}^{(n+1)} = e^{\text{obs}} + [\bar{e}^{(n)} - e^{(n)}]. \quad (27)$$

This is the same as equation (22) except for the way in which the terms are grouped.

Although the AIM-DS inversion would generally begin with the application of  $\bar{F}^{-1}$  to the observations as suggested by equation (17), it is not necessary to do this. Suppose we wished to start with an arbitrary model  $m^*$ . The iteration procedure can begin by applying a forward mapping  $\bar{F}$  which has the property that  $\bar{F}\bar{F}^{-1} = \mathcal{J}_e$ . It is convenient to

refer to  $\bar{F}$  as an approximate forward mapping. The first estimate of the corrected data is  $\bar{e}^{(1)} = \bar{F}(m^*)$  and the iteration formula (27) can be subsequently used to complete the inversion.

The introduction of a pair of mappings  $\bar{F}$  and  $\bar{F}^{-1}$  which have the property that  $\bar{F}\bar{F}^{-1} = \mathcal{J}_e$  is very useful in the AIM-DS formulation. The reasons are two-fold. Firstly, in some non-linear problems it is possible to derive an explicit form for  $\bar{F}$  perhaps based upon asymptotic analysis or from an understanding of the primary physics. This, plus the requirement that  $\bar{F}\bar{F}^{-1} = \mathcal{J}_e$ , can be used to define  $\bar{F}^{-1}$ . We shall present an example of this in the parametric inverse problem given in the next section. The second area of importance can be seen by writing equation (27) in terms of the approximate forward mapping. Substituting  $\bar{e} = \bar{F}(m)$  yields

$$\bar{F}[m^{(n+1)}] = e^{\text{obs}} + \bar{F}[m^{(n)}] - F[m^{(n)}]. \quad (28)$$

We note that if  $\bar{F}$  is a linear operator, then an updated model can be found by solving a linear inverse problem. It is important to note that the assumption that  $\bar{F}$  is linear *does not* imply that  $\bar{F}^{-1}$  is also linear. In practice, the linear inverse problem associated with solving equation (28) is usually ill-posed and requires some form of regularization. Regularization generally involves the introduction of an objective function defined in terms of (semi) norms on the spaces of models and data, and a regularization parameter controlling the trade-off between misfit and model norm. This regularization introduces non-linearity into  $\bar{F}^{-1}$  if any of the following conditions exist: the regularization parameter is not held fixed; if a semi-norm is used to define the objective function; if a norm other than an  $L_2$  norm is used. There are special circumstances when  $\bar{F}^{-1}$  is linear, for example, when one is working in a Hilbert space with fixed regularization parameter, however this circumstance is rarely of interest in practical problems. The rare case when  $\bar{F}^{-1}$  is linear is trivial.

### 3 AIM INVERSION: PARAMETRIC EXAMPLE

A simple example which illustrates the essential features of the AIM inversion can be formulated by considering the electromagnetic induction caused by a periodic uniform horizontal magnetic field incident upon a conducting earth. The incident field has an angular frequency  $\omega$  and time dependence  $e^{i\omega t}$ . The electrical structure of the earth is a function of depth only and is specified by the conductance or vertically integrated conductivity.

Under general circumstances the conductance function  $\tau(z)$  is sought from a finite number of surface measurements. Here we consider only simple structures which have a conductance sheet of magnitude  $\tau_1$  at the surface and a perfect conductor at depth  $z = h$ . Thus

$$\tau(z) = \begin{cases} \tau_1 & \text{if } z < h, \\ \infty & \text{if } z \geq h. \end{cases} \quad (29)$$

Effectively the conductivity is completely represented by a two parameter set  $(\tau_1, h)$ . We denote the space of two parameter conductance models by  $\mathcal{H}_{m_2}$  and denote the  $F$

image of  $H_{m_2}$  by  $H_{e_2}$ . We select as data, apparent resistivity as a function of frequency.

The exact forward mapping  $F$  can be derived from Maxwell's equations in 1-D. Following Parker (1980), the magnetic field induces a perpendicular, horizontal electric field  $E(z, \omega)$  governed by the integral equation

$$E(z, \omega) = E(h, \omega) - (h - z)E'(h, \omega) - i\omega\mu_0 \int_z^h (z - y)E(y, \omega) \tau(y), \quad (30)$$

where the prime denotes differentiation with respect to  $z$ . The apparent resistivity is defined by

$$\rho_A(\omega) = \mu_0\omega \left| \frac{E(0, \omega)}{\partial E(z, \omega)/\partial z|_{z=0}} \right|^2. \quad (31)$$

In the case where  $\tau(z) \in H_{m_2}$  this reduces to

$$\rho_A(\omega) = \frac{\mu_0\omega h^2}{1 + (\mu_0\omega h\tau_1)^2}. \quad (32)$$

Equation (32) defines the exact forward mapping for any  $\tau(z) \in H_{m_2}$ .

There are two asymptotic limits to the exact forward mapping given in equation (32). The low- and high-frequency limits are

$$\lim_{\omega \rightarrow 0} \rho_A(\omega) = \mu_0\omega h^2, \quad \lim_{\omega \rightarrow \infty} \rho_A(\omega) = \frac{1}{\mu_0\omega\tau_1}. \quad (33)$$

These are straight lines when plotted on a double logarithmic plot. The intersection of the asymptotes occurs at a central frequency  $\omega_c = (\mu_0 h \tau_1)^{-1}$ . The primary physics behind these relations is that an electromagnetic wave attenuates as  $\omega^{-1/2}$  as it propagates in a conductor. Consequently the high-frequency limit corresponds to the electromagnetic wave attenuating completely in the surface conducting sheet and being insensitive to the perfect conductor at depth  $h$ . The low-frequency limit corresponds to the electromagnetic wave propagating to the perfect conductor; it is not attenuated by the surface sheet and it only provides information about the depth to the perfect conductor.

To illustrate the exact forward and asymptotic mappings we consider a numerical example with  $\tau = 20$  S and  $h = 500$  m. Fig. 3(a) shows the model and Fig. 3(b) shows the resultant true data and the asymptotic curves. The data and the asymptotic curves are in reasonable agreement provided that the correct asymptotic branch is used, that is, if the low-frequency asymptote is used when  $\omega < \omega_c$ , and if the high-frequency asymptote is used when  $\omega > \omega_c$ . The central frequency for this example is  $\omega_c = 79.58$  rad s $^{-1}$  or, in terms of linear frequency,  $f_c = 12.67$  Hz. The largest discrepancies occur at  $\omega = \omega_c$ , where apparent resistivities differ by about a factor of 2. This agreement, and the fact that the asymptotic mappings incorporate the primary physics, leads us to choose the asymptotic formulae to construct both approximate forward and inverse mappings.

The approximate forward mapping is defined as follows. Given model parameters  $(\tau_1, h)$ , the apparent resistivity is

estimated as

$$\rho_A(\omega) = \begin{cases} \mu_0\omega h^2, & \omega \leq \omega_c, \\ \frac{1}{\mu_0\omega\tau_1^2}, & \omega \geq \omega_c. \end{cases} \quad (34)$$

The approximate inverse mapping is defined to be some fitting procedure from a number  $N$  of coordinate pairs  $(\rho_{A_i}, \omega_i)$ ,  $i = 1, \dots, N$  from data space to yield a two parameter model  $m \in H_{m_2}$ . Since we are looking only for two parameters it is sufficient to have  $N = 2$ . We select two pairs  $(\rho_{A_1}, \omega_1)$  and  $(\rho_{A_2}, \omega_2)$  where  $\omega_1 < \omega_2$ . The approximate inverse mapping is defined as

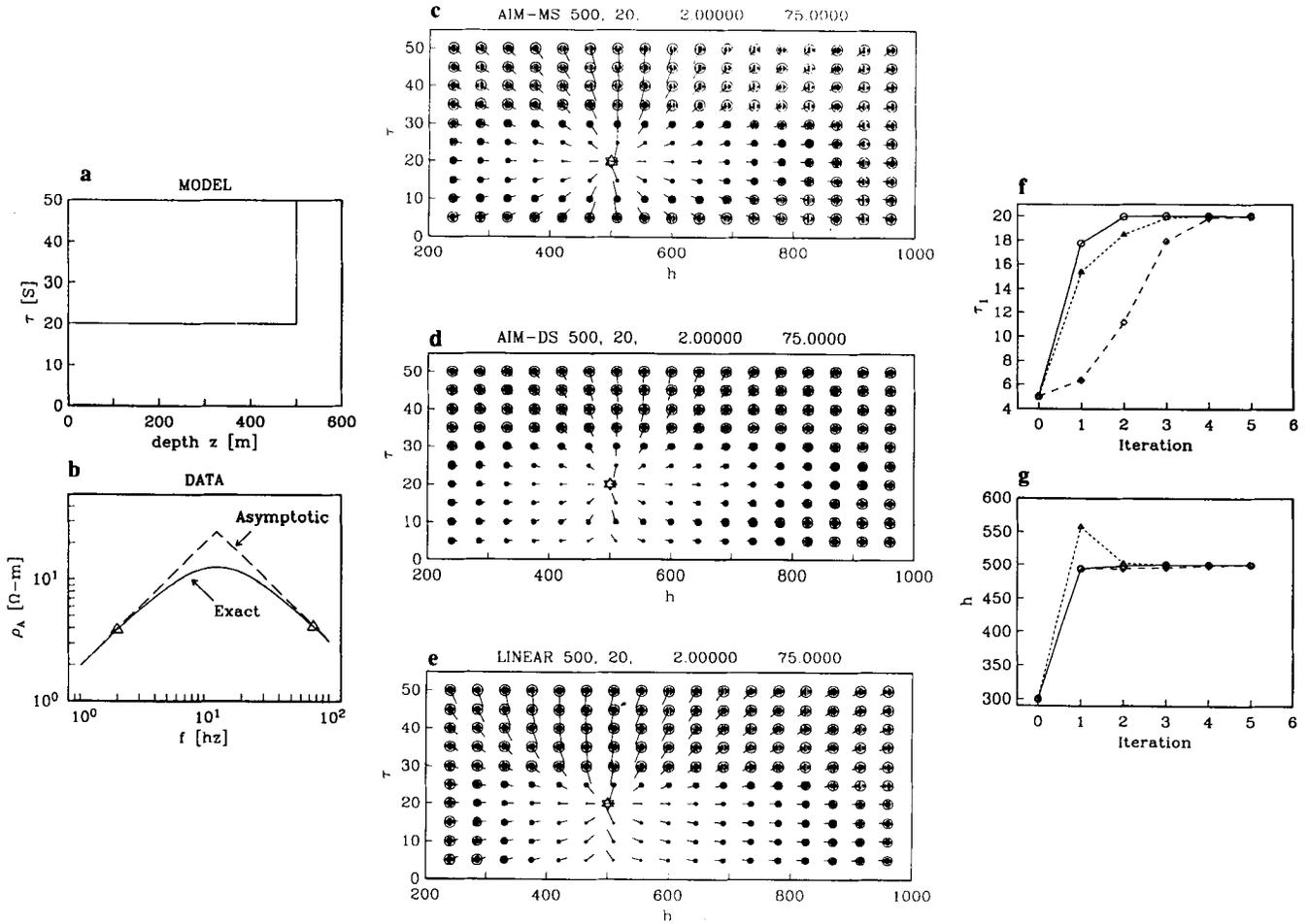
$$h = \sqrt{\rho_{A_1}/\mu_0\omega_1}, \quad \omega_1 \leq \omega_c, \\ \tau_1 = 1/\sqrt{\rho_{A_2}\mu_0\omega_2}, \quad \omega_2 \geq \omega_c. \quad (35)$$

In the derivation of the approximate inverse mapping we have assigned the low-frequency pair  $(\rho_{A_1}, \omega_1)$  to the low-frequency asymptote and thus that datum yields the depth to the perfect conductor. The high-frequency asymptote yields the surface conductance. For simplicity we assume *a priori* knowledge that  $\omega_1 < \omega_c < \omega_2$ . In fact  $\omega_c$  is easily estimated from  $\rho_A(\omega)$  since  $\rho_A(\omega)$  has a maximum at  $\omega_c$ .

The exact forward mapping, equation (32), and the approximate inverse mapping, equation (35), can be used to invert the data with the AIM technique. We first demonstrate the AIM-MS inversion and begin by choosing a data element  $e^{\text{obs}} \in H_{e_2}$  to be inverted. Each data element consists of two apparent resistivities; we choose linear frequencies  $f_1 = 2$  Hz and  $f_2 = 75$  Hz. These are shown in Fig. 3(b) as triangles. Following the procedure outlined in the Section 2 and defined in equation (15) we form a series of models  $\{m_i\}$  each of which is represented as a coordinate pair  $(\tau_1, h)$ . From the development of equation (35) it is expected that the approximate inverse mapping would rapidly give an excellent representation of the true model. Table 1(a) gives the predicted model and predicted data for successive iterations. Convergence is rapid and only three iterations are required before the model parameters are recovered to five significant figures.

The AIM-MS inversion given above began with a model obtained by applying the approximate inverse mapping directly to the data. This is not necessary as equation (15) indicates. To illustrate this we show in Fig. 3(c) the direction and magnitude of the perturbation step generated by the inversion algorithm starting with any model  $(\tau_1, h)$ ,  $5 < \tau_1 < 55$ ,  $100 < h < 1000$ . In Fig. 3(c) a small pointer is used to indicate the perturbation step direction and a circle of varying radius indicates the magnitude of the perturbation step; the larger the radius, the greater the perturbation step. This diagram indicates that good convergence is expected for a large range of starting models. As a specific example we show in Fig. 3(f, g) (solid curve) the convergence when the starting model is  $(\tau_1, h) = (5, 300)$ .

We next demonstrate the AIM-DS inversion. The series of models generated by the AIM-DS algorithm, equation (27), are given in Table 1(b). Convergence is rapid for this example where the first model was obtained by applying the approximate inverse mapping to the observations. Fig. 3(d)



**Figure 3.** A two parameter model consisting of a 20 S surface conductance sheet and a perfectly conducting basement at 500 m is shown in (a). The apparent resistivity versus frequency data space response generated from that model is shown in (b). Data obtained using the exact forward mapping are shown as the solid line and data generated from the approximate forward mapping is shown as the dashed line. The two frequencies used in the parametric inversion are denoted by  $\Delta$ 's. The perturbation steps for AIM-MS, AIM-DS, and linearized methods beginning at different locations in model space are shown in (c)–(e) respectively. For each point in the  $(\tau_1, h)$  plane these plots show the magnitude and direction of the perturbation step predicted by the inversion algorithm. The magnitude of the perturbation step is represented by the radius of the hatched circle, the direction is indicated by the pointer. Specific convergence paths for the parametric model inversions for a starting model  $(\tau_1, h) = (5, 300)$  are shown in panels (f) and (g). The predicted surface conductance at each iteration is given in (f) and the depth to the perfect conductor is given in (g). The results for the AIM-MS, AIM-DS and linearized inversions are marked by  $\circ$ ,  $\diamond$ ,  $\Delta$ , respectively.

shows the perturbation steps for starting models in the domain  $(\tau_1, h)$ ,  $5 < \tau_1 < 55$ ,  $100 < h < 1000$  and is the AIM-DS analogue of Fig. 3(c). Good convergence is indicated for all models. Comparison between Fig. 3(c) and Fig. 3(d) clearly shows that AIM-MS has faster convergence than AIM-DS for small conductances. This is also emphasized in Fig. 3(f) (course dash) where convergence is shown for a starting model  $(\tau_1, h) = (5, 300)$

It is of interest to compare the AIM results using  $\tilde{F}^{-1}$  based on the asymptotic mappings of equation (35) with those obtained by an  $\tilde{F}^{-1}$  based on conventional linearization. To find an  $\tilde{F}^{-1}$  based on conventional linearization a  $2 \times 2$  sensitivity matrix can be generated by performing a Taylor expansion of equation (32) and applying it to the two data to be inverted. Applying the inverse of this matrix to the discrepancy between the observed and predicted data yields the model perturbation.

Note that this linearization is an example of the rare case where  $\tilde{F}^{-1}$  is linear: AIM-MS and AIM-DS then degenerate to the same algorithm,  $\delta m = \tilde{F}^{-1} \delta e$ . This rare degeneracy happens only because the example has the same number of parameters as exact data. In Table 1(c) we show the series of models obtained by the AIM inversion based on linearization. We refer to this inversion as the linearized inversion. The initial starting model for the linearized solution was the same as that in the AIM inversions in Table 1(a, b). To obtain an overview of the convergence properties we show in Fig. 3(e) the first perturbation step for the linearized inversion initiated from a variety of starting models. This plot is analogous to those in Fig. 3(c, d). A specific convergence plot for a starting model  $(\tau_1, h) = (5, 300)$  is given in Fig. 3(f, g). Comparison of Figs 3(c), (d) and (e) indicates that the linearized inversion shows better convergence than the AIM-DS inversion,

**Table 1.** The recovered surface conductance and depth to the perfect conductor for the parametric example are given as a function of iteration. The predicted apparent resistivities at frequencies  $f_1 = 2$  Hz and  $f_2 = 75$  Hz are shown in the last two columns. Tables 1(a), (b), and (c) respectively show the results for the AIM-MS, AIM-DS, and linearized inversion.

## a AIM-MS Inversion

Iteration	Conductance	Depth h	Apparent resistivity 1	Apparent resistivity 2
0	493.88	20.283	3.7578	3.9912
1	499.95	19.997	3.8510	4.1058
2	500.00	20.000	3.8517	4.1047
3	500.00	20.000	3.8518	4.1047
4	500.00	20.000	3.8518	4.1047

## b AIM-DS Inversion

Iteration	Conductance	Depth h	Apparent resistivity 1	Apparent resistivity 2
0	493.88	20.283	3.7578	3.9912
1	499.87	20.009	3.8498	4.1012
2	500.00	20.000	3.8518	4.1045
3	500.00	20.000	3.8518	4.1047
4	500.00	20.000	3.8518	4.1047

## c Linearized Inversion

Iteration	Conductance	Depth h	Apparent resistivity 1	Apparent resistivity 2
0	493.88	20.283	3.7578	3.9912
1	500.04	19.994	3.8524	4.1071
2	500.00	20.000	3.8518	4.1047
3	500.00	20.000	3.8518	4.1047
4	500.00	20.000	3.8518	4.1047

however, worse convergence than the AIM-MS inversion. This is most evident in the region  $\tau < 20$  and is emphasized in Fig. 3(f) (fine dash). Note in particular that the linearized inversion overshoots the desired depth at the first iteration contrary to the behaviour of the AIM-DS and AIM-MS algorithms.

This simple parametric example using AIM-MS, AIM-DS and linearized inversions demonstrates some important features of the approximate mapping inversion technique. First, it shows that the technique can converge rapidly when the approximate mapping is a good approximation to the exact mapping. This convergence seems to be maintained even when the approximate inverse mapping is a poorer approximation to the exact mapping. For example, when the above inversions used frequencies close to the centre frequency, thereby accentuating the difference between the true and approximate forward mappings, both AIM algorithms converged to the desired solution. Convergence was also noted even under more extreme cases where both frequencies were less than the centre frequency or both were larger. A second feature of the AIM algorithms is that the convergence paths can be significantly different for

different approximate inverse mappings: the AIM-MS and AIM-DS inversion based on asymptotic mappings (35) have very different convergence paths compared with the AIM inversion based on linearization. This follows from the fact that the approximate physics used to generate the approximate mapping may be non-linear and consequently the approximate inverse mapping is also non-linear. As such, an AIM algorithm with appropriate  $\tilde{F}^{-1}$  has the potential for superior convergence compared to conventional linearized inversion algorithms.

## 4 AIM INVERSION: MT IN 1-D

The parametric examples illustrate the essential steps in the AIM inversion and they also provide some indication about the robustness and convergence properties of the algorithms. The examples are restrictive however, in the sense that model space and data space are each two parameter subspaces. Our ultimate goal is to use the AIM inverse formalism to solve large-scale geophysical problems. Therefore, as a next example, we consider the inversion of many MT data to recover a conductivity  $\sigma(z)$ . To keep a formulation which is analogous to the parametric example we choose apparent resistivities as data. The approximate forward and inverse mappings are again derived through asymptotic analysis. These mappings have been investigated by many workers, (e.g. Niblett & Sayn-Wittgenstein 1960; Word, Smith & Bostick 1970; Bostick 1977); a review can be found in Whittall & Oldenburg (1991). Although the mappings vary in particulars, they are all derived from the physical understanding that the attenuation of an electromagnetic wave is governed by the conductance of the medium.

For this example we take the following asymptotic mappings. For the forward mapping

$$\tilde{F}: \quad \rho_A = \frac{z}{\xi \tau}, \quad \omega = \frac{\xi}{\mu_0 z \tau}, \quad (36)$$

and for the inverse mapping

$$\tilde{F}^{-1}: \quad \tau = \sqrt{\frac{1}{\mu_0 \omega \rho_A}}, \quad z = \xi \sqrt{\frac{\rho_A}{\mu_0 \omega}}. \quad (37)$$

We have chosen to include a stretching factor  $\xi$  in depth. This is a slight generalization over the usual equations although depth stretchings have been incorporated in the past; e.g. Gamble (1983). The incorporation of  $\xi$  will produce a different model for the AIM-MS solution than was given in Whittall & Oldenburg (1991) in their summary of the work of Goldberg *et al.* (1982).

The asymptotic mappings establish a bijection between a 2-D data space  $(\rho_A, \omega)$  and a 2-D model space  $(\tau, z)$ . For example the data space function  $\rho_A(\omega) = \omega^{-1}$  maps to the model space function  $\tau(z) = \sqrt{\mu_0}$ . Note that these approximate mappings are a more general form of equations (34) and (35) used in the parametric example.

There are a number of ways to implement the above mappings. Here we choose to recover a layered conductivity structure  $(\sigma_j, z_j)$ ,  $j = 1, \dots, N$  where  $z_j$  is the depth to the bottom of the  $j$ th layer of conductivity  $\sigma_j$ . We also define  $\tau_j$  to be the conductance at depth  $z_j$ , that is,  $\tau_j$  is the integrated conductivity from the surface to the depth  $z_j$ . Given a set of

data  $(\rho_{\Lambda_j}, \omega_j)$ ,  $j = 1, \dots, N$  the layered conductivity structure is given by:

$$\begin{aligned}
 h_1 &= z_1, & \sigma_1 &= 1/\rho_{\Lambda_1}, \\
 z_j &= \xi \sqrt{\frac{\rho_{\Lambda_j}}{\mu_0 \omega_j}}, & \tau_j &= \sqrt{\frac{1}{\mu_0 \omega_j \rho_{\Lambda_j}}}, \\
 h_j &= z_j - z_{j-1}, & \sigma_j &= (\tau_j - \tau_{j-1})/h_j.
 \end{aligned}
 \tag{38}$$

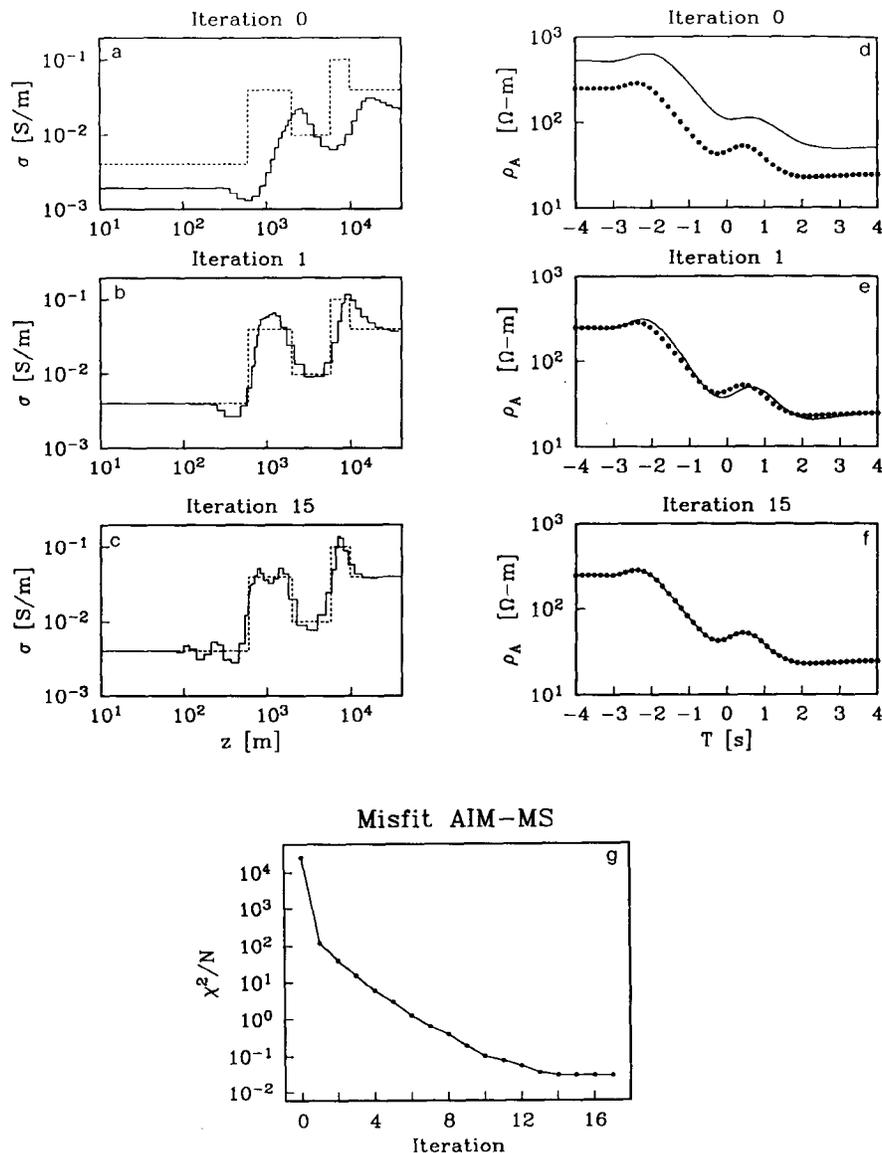
In the inverse problem flexibility exists in defining the model and the data. Ultimately the set  $(\sigma_j, z_j)$ ,  $j = 1, \dots, N$  is the physical quantity desired and this is considered as the 'model' when the forward mapping is carried out. However, conductivity in earth materials varies over many orders of magnitude. With respect to the inverse problem, it is better to consider  $(\ln \sigma, \ln z)$  as the model. The AIM-MS iterative

equation (15) then becomes

$$\begin{aligned}
 \ln \sigma^{(n+1)} &= \ln \bar{\sigma}^{\text{obs}} + \ln \sigma^{(n)} - \ln \bar{\sigma}^{(n)}, \\
 \ln z^{(n+1)} &= \ln \bar{z}^{\text{obs}} + \ln z^{(n)} - \ln \bar{z}^{(n)},
 \end{aligned}
 \tag{39}$$

where  $\bar{\sigma}$  and  $\bar{z}$  are evaluated via equation (38).

To illustrate the AIM-MS inversion algorithm we use the test model shown in Fig. 4. Since each datum will generate a single layer, we have used 50 accurate data spanning a frequency range  $10^{-4}$ – $10^4$  Hz. This ensures that a reasonably complicated model can be generated by the inversion. Models at selected iterations, and fits to the data are shown in Fig. 4(a–f). The models predicted by the inversion,  $\sigma^{(n)}$ , are shown plotted (solid curve) over the test model (dashed) which was used to generate the data to be inverted,  $e^{\text{obs}}$ . Opposite each model plot is the predicted

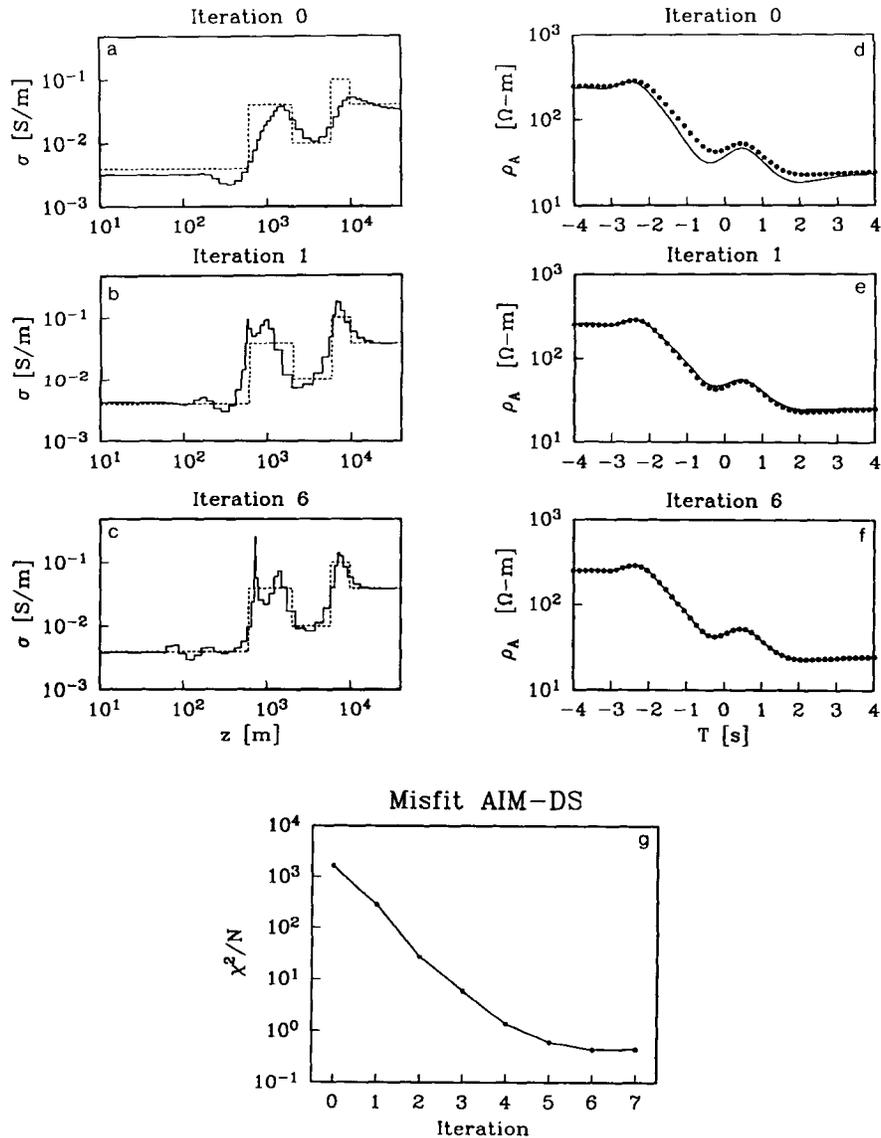


**Figure 4.** The convergence of the AIM-MS algorithm using the 1-D asymptotic mappings. Shown on the left are the true model (dashed) and the predicted model (solid) at iterations 0, 1 and 15. Shown on the right are the true data (\*) and the predicted data (solid). The normalized  $\chi^2$  misfit as a function of iteration (evaluated by assuming that the data are contaminated with 1 per cent Gaussian error) is shown at the bottom.

data,  $e^{(n)}$  (solid curve) and  $e^{obs}$  (denoted by \*). A stretch factor of  $\xi = 2.1$  was selected for this example. The model at iteration 0 (Fig. 4a) is the result of applying the approximate inverse mapping directly to the observations. The conductivity is greatly underestimated compared to the true model and the computed apparent resistivities are biased upward compared to the true data in Fig. 4(d). However, a single correction using the model space mapping error produces the results at iteration 1. Dramatic improvements in the model (Fig. 4b) and in the fit to the data (Fig. 4e) are evident; a single non-linear perturbation step has reduced the misfit by more than two orders of magnitude. To quantify the misfit we have initially assigned a Gaussian error of 1 per cent to each amplitude and normalized the resultant  $\chi^2$  misfit by the number of data. The normalized misfit,  $\chi^2_N = \chi^2/N$ , at successive iterations is shown in Fig. 4(g). We note a monotonic decrease in the misfit and by the 15th iteration the data are fit to within  $\chi^2_N = 0.03$ .

Equivalently, the rms fit to the data is 0.17 per cent. The final model (Fig. 4c) is a good representation of the true model but it does exhibit oscillations which may be undesirable.

Next, the AIM-DS algorithm is implemented with the same approximate inverse mapping defined by equations (37) with a stretch factor  $\xi = 1.2$ . Models at selected iterations and the misfit convergence are shown in Fig. 5(a-f). The  $\chi^2_N$  misfit monotonically decreases by about an order of magnitude for each of the first four iterations and then plateaus at a value of about  $\chi^2_N = 0.44$ . At this final stage each datum is fit to within about 0.66 per cent. We note the increase in roughness of the model at progressive iterations. This results because the final data to be inverted are obtained through a sequence of additive perturbations and hence data having significant structure are eventually generated. If there is no regularization in the approximate inverse mapping then this increase in structure on the



**Figure 5.** The convergence of the AIM-DS algorithm using the 1-D asymptotic mappings. Shown on the left are the true model (dashed) and the predicted model (solid) at iterations 0, 1 and 6. Shown on the right are the true data (\*) and the predicted data (solid). The normalized  $\chi^2$  misfit as a function of iteration (evaluated by assuming that the data are contaminated with 1 per cent Gaussian error) is shown at the bottom.

modified data will be transformed to structure on the recovered model. We note for algorithms of this type that the reciprocal of the number of iterations acts as a regularization parameter (Defrise & De Mol 1987), thus increasing structure with increasing iteration number is to be expected.

We are not proposing that the 1-D MT inverse problem be solved with an AIM inversion which uses an asymptotic inverse as an approximate inverse mapping and employs no explicit regularization. The purpose of the above illustrations is to show that the AIM formalism, when applied in this manner, can converge to a model that (at least approximately) reproduces the data. The steady decrease in misfit is extremely encouraging. On the other hand, the progressive increase in model roughness is an indication that the algorithm is not robust, at least if arbitrarily small misfits are required. This negative aspect of the constructed model follows from the fact that the asymptotic mappings, by themselves, are a poor choice for an approximate inverse mapping. There is no allowance for errors on the data, there is no attempt within the algorithm to minimize an objective function of the desired model, and there is no way to incorporate additional information about the conductivity structure. The rather unattractive oscillations of the conductivity structure are a reflection of these omissions. They do not however, detract from the fact that the algorithms have marched progressively toward a solution.

Building upon these observations we now choose to design an approximate inverse mapping that generates a model which minimizes an objective function and incorporates data errors directly. The AIM-DS formulation is adopted and specifically, we use equation (28) in which an approximate forward mapping  $\tilde{F}$  is introduced. If  $\tilde{F}$  is a linear mapping then it is straightforward to find an updated model  $m^{(n+1)}$  by solving a linear inverse problem. In doing so, the linear inverse can be constructed to minimize an objective function of the model, subject to appropriate misfit of the data and any additional linear constraints that might be available on the model; this effectively defines  $\tilde{F}^{-1}$ .

The implementation of the AIM-DS procedure given by equation (28) requires the specification of  $\tilde{F}$ . One possibility for  $\tilde{F}$  can be obtained by linearizing the exact forward mapping  $F$  about a model  $m^{(n)}$ . Applied to an arbitrary model  $m$  this yields

$$F(m; \omega) = F[m^{(n)}; \omega] + \int g[m^{(n)}; z, \omega][m(z) - m^{(n)}(z)] dz + \dots, \quad (40)$$

where  $g$  is the associated kernel function (Oldenburg 1979) and the ellipsis represents higher order terms in the functional expansion. An approximate forward mapping can be obtained by keeping terms up to first order. We define

$$\tilde{F}(m; \omega) = F[m^{(n)}; \omega] - \int g[m^{(n)}; z, \omega]m^{(n)}(z) dz + \int g[m^{(n)}; z, \omega]m(z) dz. \quad (41)$$

By writing the first two terms as  $\Delta^{(n)}$ , the approximate

forward mapping at the  $n$ th iteration is given by

$$\tilde{F}(m; \omega) = \Delta^{(n)} + \int g[m^{(n)}; z, \omega]m(z) dz. \quad (42)$$

Substitution of equation (42) into equation (28) yields

$$\int g[m^{(n)}; \omega]m^{(n+1)}(z) dz = e^{\text{obs}}(\omega) - e^{(n)}(\omega) + \int g[m^{(n)}; \omega]m^{(n)}(z) dz. \quad (43)$$

It is noted that the constant  $\Delta^{(n)}$  does not appear in the final equation. For notational purposes and for application, the approximate forward mapping could equivalently have been defined by omitting  $\Delta^{(n)}$  in equation (42).

The inversion proceeds by minimizing an objective function which includes data misfit and model norm. It is noted that this special choice of the AIM-DS algorithm has been successfully used in the past. Oldenburg (1983) used it to find upper and lower bounds to average values of the conductivity in a 1-D MT inversion. Also, Constable, Parker & Constable (1987), Smith & Booker (1988a), and Dosso & Oldenburg (1989) have used this formulation to obtain minimum structure conductivity models for the 1-D MT inverse problem.

We now illustrate this form of AIM-DS applied to the 1-D MT inverse problem. In defining the inversion algorithm, there are choices as to definition of the model for the inverse problem, characterization of the data, specification of the objective function to be minimized, and method of regularization. In the forward mapping the parameter of interest is conductivity, however  $\ln(\sigma)$  will be used as the 'model' for the inverse problem. Let us denote  $(m_1, m_2, \dots, m_{n_z})$  to be the discretized array of  $\ln(\sigma)$  values associated with a set of  $n_z$  depth partitions. There is also a choice for what are regarded as 'data' within the 1-D inversion. Here we choose the complex variable

$$e_l = \frac{\partial E(z, \omega_l) / \partial z |_{z=0}}{\omega_l E(0, \omega_l)}, \quad l = 1, \dots, n_f, \quad (44)$$

specified at  $n_f$  frequencies. At each frequency the complex datum is represented as an amplitude and phase and so  $2n_f$  real data are used in the inversion.

With the above definitions, equation (43) is written as

$$\sum_i^{n_z} A_{ik} m_i^{(n+1)} = e_k^{\text{obs}} - e_k^{(n)} + \sum_i^{n_z} A_{ik} m_i^{(n)}, \quad k = 1, \dots, 2n_f, \quad (45)$$

where  $A_{ik}$  is the integral of the  $k$ th kernel function over the  $i$ th depth partition.

We minimize an objective function

$$\phi = \sum_k^{2n_f} \left| \frac{e_k^{\text{obs}} - e_k^{(n)}}{\epsilon_k} \right| + \beta \sum_i^{n_z-1} |m_{i+1}^{(n+1)} - m_i^{(n+1)}|, \quad (46)$$

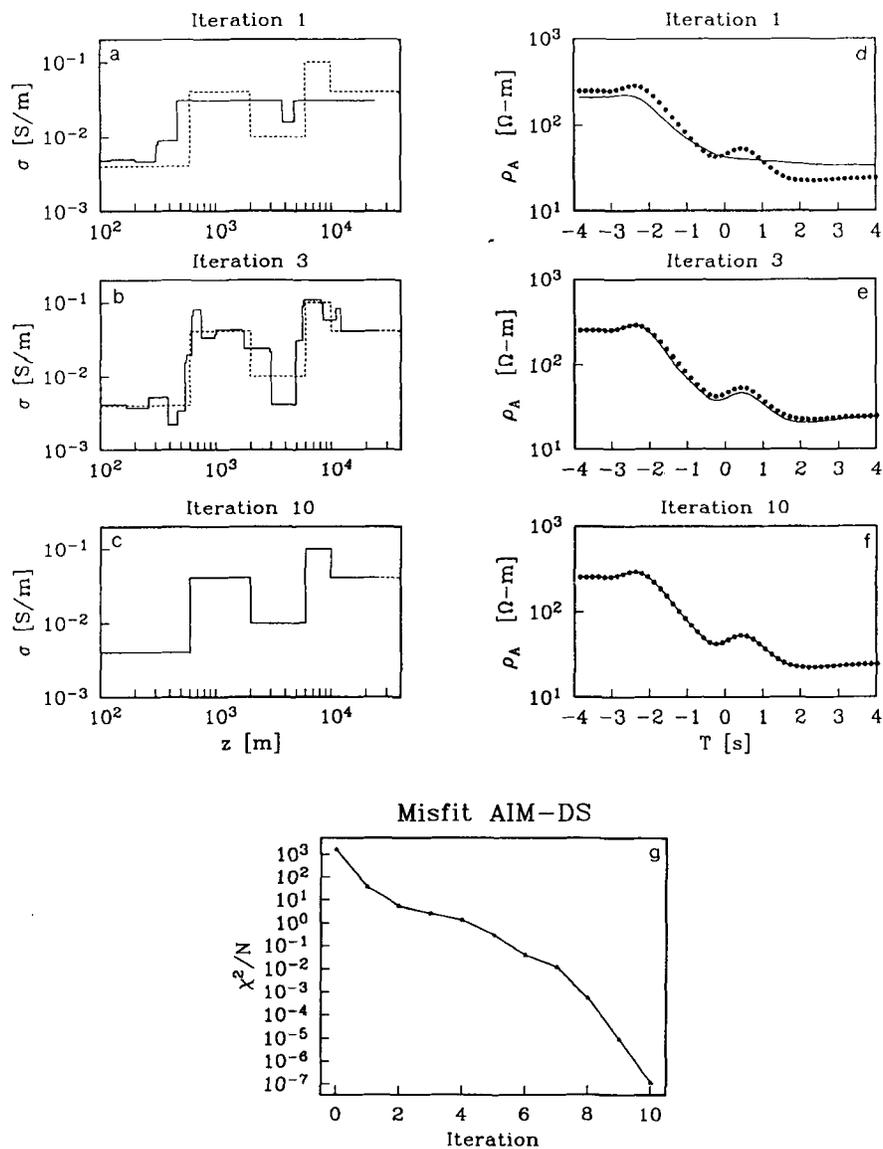
where  $\epsilon_k$  is the standard deviation of the  $k$ th datum. The frequencies associated with the data are the same as in the preceding example, so  $n_f = 50$ . However, here we choose to parametrize the depth by  $n_z = 169$  logarithmically spaced depth partitions; these remain fixed throughout the

inversion. The parameter  $\beta$ , held constant throughout the inversion, controls the trade-off between data misfit and model variation. The problem of minimizing (46) subject to the constraints in (45) is solved using linear programming (LP) techniques.

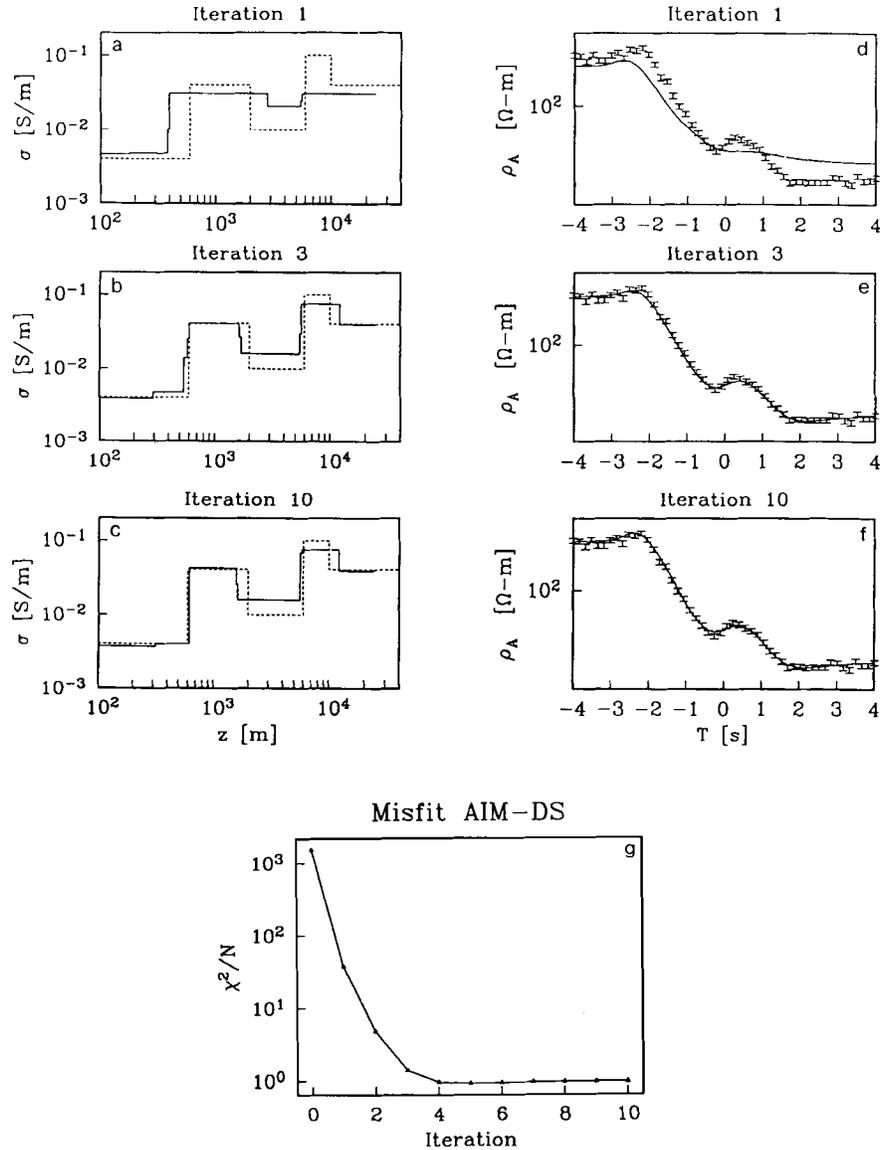
To illustrate the technique we have inverted two sets of data: the exact data used in the previous examples, and data with Gaussian independent noise of standard deviations 5 per cent added to the apparent resistivities and  $2^\circ$  to the phases. The result of the inversion of the exact data is shown in Fig. 6(a-f). The true model is shown dashed and the corresponding data to be inverted,  $e^{\text{obs}}$ , as \*'s. The starting model is a half-space. The models  $m^{(n)}$  produced from the inversion, are shown for iterations  $n = 1, 3, 10$  and are plotted as solid curves. The predicted data  $e^{(n)}$  are also shown. After 10 iterations the inversion result and the true

model are indistinguishable. Of course the MT inverse problem, even with data at 50 frequencies, is non-unique and there are other models (some very different) which will reproduce these data to within the same accuracy. The fact that we have such remarkable correspondence between the constructed and true models is an excellent illustration of the results that can be achieved by minimizing the 'right' objective function in an inverse problem which is inherently underdetermined. In Fig. 6(g) we show the  $\chi^2$  misfit of the predicted data  $e^{(n)}$  compared with the observed data  $e^{\text{obs}}$ .

Similar plots are shown for the inversion of data with noise in Fig. 7. In this case the data to be inverted are plotted with error bars. After 10 iterations a 'minimum structure model' is obtained with a misfit  $\chi_N^2 = 0.98$  for the apparent resistivities and  $\chi_N^2 = 1.1$  for phases. In the past much theoretical effort has been devoted to finding the



**Figure 6.** The convergence of the AIM-DS algorithm using the 1-D linearized mapping for data without noise. Shown on the left are the true model (dashed) and the predicted model (solid) at iterations 1, 3 and 10. Shown on the right are the true data (\*) and the predicted data (solid). The normalized  $\chi^2$  misfit as a function of iteration (evaluated by assuming that the standard deviations are 5 per cent on the apparent resistivities and  $2^\circ$  on the phases) is shown at the bottom.



**Figure 7.** The convergence of the AIM-DS algorithm using the 1-D linearized mapping for data contaminated with Gaussian noise of standard deviations 5 per cent on apparent resistivities and  $2^\circ$  on the phases. Shown on the left are the true model (dashed) and the predicted model (solid) at iterations 1, 3 and 10. Shown on the right are the true data with error bars and the predicted data (solid). The normalized  $\chi^2$  misfit as a function of iteration is shown at the bottom.

proper choice of regularization parameter,  $\beta$  such that the appropriate data misfit is achieved. For the inversions considered here it suffices to find  $\beta$  by trial and error.

The preceding 1-D inversion will form the basis of the subsequent 2-D inversion so it is appropriate at this time to mention some characteristics of the inversion algorithm. We have used a LP package (Marsten 1981) designed to take advantage of a sparse tableau in order to minimize memory requirements. The user has some control over partial pricing and the candidate list which are required to avoid spurious degeneracy. The package also has the capability of starting with a predefined basis. We start a new model iteration with the final LP basis from the preceding model iteration. This can yield a decrease in CPU requirements of orders of magnitude, particularly when small changes are made to the models at each iteration. For the inversion of data with

noise shown above  $n_f = 50$ ,  $n_z = 169$ , the initial iterations took  $\sim 200$  s, the final iterations took  $\sim 5$  s, with an average CPU time per iteration for 10 iterations of  $\sim 70$  s. All computations presented in this paper were performed on a SUN 4/370 with 8 MB of memory.

## 5 AIM INVERSION: 2-D MAGNETOTELLURICS

The motivation for the AIM inversion technique is to develop a computationally efficient method for solving large multidimensional inverse problems. In the preceding section we demonstrated that the AIM inversion could be used successfully in the 1-D inverse problem. We now demonstrate the AIM inversion in 2-D magnetotellurics.

We begin by defining the exact forward mapping which

follows from Maxwell's equations. When the conductivity is a function of two coordinates,  $\sigma = \sigma(y, z)$ , Maxwell's equations decouple. By neglecting the displacement current and assuming harmonic time dependence  $e^{i\omega t}$  for the fields, we obtain the TE mode equations in terms of the strike component of the electric field  $E_x = E_x(y, z, \omega)$ :

$$\frac{\partial^2 E_x}{\partial z^2} + \frac{\partial^2 E_x}{\partial y^2} - i\omega\mu_0\sigma E_x = 0, \quad (47)$$

$$H_y = -\frac{1}{i\omega\mu_0} \frac{\partial E_x}{\partial z}, \quad H_z = \frac{1}{i\omega\mu_0} \frac{\partial E_x}{\partial y},$$

and the TM mode equations in terms of the strike component of the magnetic field  $H_x = H_x(y, z, \omega)$ ,

$$\frac{\partial}{\partial z} \left( \frac{1}{\sigma} \frac{\partial H_x}{\partial z} \right) + \frac{\partial}{\partial y} \left( \frac{1}{\sigma} \frac{\partial H_x}{\partial y} \right) - i\omega\mu_0 H_x = 0 \quad (48)$$

$$E_y = \frac{1}{\sigma} \frac{\partial H_x}{\partial z}, \quad E_z = -\frac{1}{\sigma} \frac{\partial H_x}{\partial y}.$$

For a thorough discussion of these equations, appropriate boundary conditions and numerical solution, the reader is referred to Stratton (1941), Kauffman & Keller (1983) or Ward & Hohmann (1988).

To effect our forward mapping, the 2-D conductivity model is first cellularized with rectangular elements. The model is partitioned into  $n_y$  horizontal cells and  $n_z$  vertical cells and the 2-D conductivity  $\sigma(y, z)$  is partitioned to an  $n_z \times n_y$  array  $\sigma_{ij}$ ,  $i = 1, \dots, n_z$ ;  $j = 1, \dots, n_y$ . The thickness of the cells increases (usually logarithmically) with depth. For simplicity, lateral partitioning is uniform in the area of the survey and exponentially increasing outside the area of interest. The conductivity is assumed to be constant in each cell and we have chosen to put an observation site at the centre of each surface cell. The 2-D MT responses are computed using a transmission surface modelling code (Madden 1972).

To define the inversion algorithm we must choose the characterization of the data, specification of the objective function to be minimized, and method of regularization. For the 2-D MT inverse problem we choose  $\ln(\sigma)$  as the 'model' for the inverse problem. Let us denote  $m_{ij}$  to be the cellularized array of  $\ln(\sigma_{ij})$  values where the cellularization is the same as used for the exact forward mapping. We choose as data the amplitude and phase of the complex variable

$$e_{jl} = \frac{\partial E(y_j, z, \omega_l) / \partial z|_{z=0}}{\omega_l E(y_j, 0, \omega_l)}, \quad (49)$$

$$j = 1, \dots, n_y; \quad l = 1, \dots, n_f,$$

specified at  $n_y$  observation sites and at  $n_f$  frequencies.

We now turn to the important question of what is an appropriate  $\tilde{F}^{-1}$  for this problem. In the preceding examples for 1-D magnetotellurics the approximate mappings were based on the 1-D asymptotic relations: the conductance in the first skin depth was the controlling parameter. In the 2-D magnetotelluric problem we suggest that the primary physics at any measurement site is due to the conductivity structure directly under that site. The justification for this suggestion can be seen by examining the

sensitivity matrix,  $J$ , associated with the change in apparent resistivity (or phase) data,  $\delta\rho_A(y', \omega)$  resulting from a perturbation in conductivity,  $\delta\sigma(y, z)$ , ie.

$$J(\sigma; y, y', z, \omega) = \frac{\delta\rho_A(y', \omega)}{\delta\sigma(y, z)}. \quad (50)$$

In general,  $J$  is diagonally dominant with respect to  $y, y'$  i.e.  $J$  deviates most significantly from zero when the datum  $\rho_A(y', \omega)$  is directly above the perturbed conductivity  $\delta\sigma(y, z)$ . This will be illustrated in an example presented later. It follows from this choice of the dominant physics that an approximate inverse mapping,  $\tilde{F}^{-1}$ , for the 2-D magnetotelluric problem may be taken to be a composite of 1-D inversions. Clearly, there are many other choices for  $\tilde{F}^{-1}$  and that selected here may not be the best possible. However, it is physically well motivated, simple to implement, and builds on previous knowledge gained in the study of the 1-D magnetotelluric problem.

Our use of an approximate inverse mapping which is based upon compositing 1-D inversion results is not new. Geophysics abounds with examples. Ryu, Morrison & Ward (1972) composited 1-D inversions of em soundings to construct a conductivity profile across the Santa Clara Valley; Vozoff (1972) generated a 2-D conductivity section across South Texas by performing layered model interpretations of MT data, and in seismology, the plus-minus method of Hagedoorn (1959) is still used to estimate laterally varying thickness of the weathering layer. Smith & Booker (1988b) have recently used such a mapping in their iterative technique for constructing a 2-D conductivity profile that for MT data.

Taking the approximate inverse mapping to be a composite of 1-D inversions requires an algorithm which will invert data at each station to generate 1-D conductivity profile beneath that station. We select as a 1-D inversion the AIM-DS algorithm presented in the last section and which gave the results shown in Figs 6 and 7. The objective function in equation (45) however, involves only data misfit and vertical variation of the conductivity model. Our ultimate goal is to minimize structure in both the vertical and horizontal directions; it is therefore necessary to alter the objective function so that lateral variation is included. Unfortunately, this is not a straight forward procedure. We have taken two different approaches. The first involves the incorporation of a 'base-model' which is defined by the user. Lateral variability is evaluated as deviations from this model. The second approach minimizes directly an objective function of the model by solving a system of equations which involves all of the unknown model elements.

The inclusion of lateral variability in the objective function can be effected by introducing a base-model,  $\tilde{m}$ . The base-model may take a variety of forms depending upon what is known about the regional conductivity structure. It may be a 1-D conductivity model (that is, the same base-model is used for each station) or it may be a 2-D model that is estimated by some means. The base-model can be chosen at the outset and kept fixed for the entire inversion, or it may be updated after each iteration in the 2-D inverse problem. For now, we include the possibility that the base-model can be updated and we write the 1-D

objective function to be minimized at each station  $y_j$  as

$$\phi_j = \sum_k^{2n_f} \left| \frac{e_k^{\text{obs}} - e_k^{(l)}}{\epsilon_k} \right| + \beta \sum_i^{n_z-1} |m_{i+1,j}^{(l+1)} - m_{i,j}^{(l+1)}| + \gamma \sum_i^{n_z} w_i |m_{i,j}^{(l+1)} - \bar{m}_{i,j}^{(n)}|. \quad (51)$$

The index  $l$  has been explicitly included to distinguish between iterations in the 1-D inversion and iterations in the 2-D inversion which will be denoted by  $n$ . In equation (51)  $e_k^{(l)}$  is the predicted datum from a 1-D forward modelling and  $\epsilon_k$  is the standard deviation of the corresponding measured field datum. However the 'observed' datum  $e_k^{\text{obs}}$  is meant to be interpreted in a generic sense. Depending upon whether an AIM-MS or AIM-DS solution is sought, the 'observed' data will be the measured field observations, the accurate 2-D forward modelled data, or the measured data corrected by the estimated mapping error. The parameters  $\beta$  and  $\gamma$  control the trade-off between data misfit, vertical model variation, and horizontal model variation. The weighting  $w_i$  is used to control the amount of lateral variation as a function of depth and cell thickness. The approximate inverse mapping for the 2-D problem is defined by solving the above 1-D inverse problem at each station and combining the resultant conductivities into a 2-D model.

The introduction of a base-model from which model variation is measured is sometimes undesirable. This is particularly true in those circumstances where little is known about the structure under investigation and hence the base-model is merely guessed at. Since different base-models will yield different final solutions, a poor guess may predetermine a poor answer from the inversion algorithm.\* Under such circumstances we prefer to minimize an objective function which includes variation in both spatial directions and does not include a base-model. The drawback to this approach is that a larger matrix system of equations needs to be solved. Fortunately the matrix is sparse and hence LP methods can still be used.

Our formulation is based upon the use of an AIM-DS algorithm in which a 2-D linear approximate forward mapping is introduced. Since  $\bar{F}$  is a linear operator it can be used to generate an approximate inverse mapping,  $\bar{F}^{-1}$ . A regularization term based on a 2-D objective function of the model will produce the desired result. The difficulty is in the specification of  $\bar{F}$ . One possibility is suggested by the development of the 1-D AIM-DS equations (40–43). Let us consider the inversion algorithm for a particular model offset  $y = y_0$ . We begin by linearizing, in the  $z$  direction, the exact forward mapping,  $F$ , about a model  $m^{(n)}$  which in this case represents  $\sigma(y, z)$ . This yields

$$F(m; \omega)|_{y=y_0} = F[m^{(n)}; \omega]|_{y=y_0} + \int g_{1d}[m^{(n)}; y_0, z, \omega] \times [m(y_0, z) - m^{(n)}(y_0, z)] dz + \dots, \quad (52)$$

where  $g_{1d}$  is the 1-D kernel function associated with the conductivity  $\sigma(y, z)$  at model offset  $y_0$ . Defining the

\* In an attempt to overcome some of these problems, an extensive investigation into the use of base-models with 1-D inversions interpolated for the 2-D MT inverse problem has been made by Smith (1988).

approximate forward mapping by keeping only the first two terms yields

$$\bar{F}(m; \omega)|_{y=y_0} = F[m^{(n)}; \omega]|_{y=y_0} + \int g_{1d}[m^{(n)}; y_0, z, \omega] [m(y_0, z) - m^{(n)}(y_0, z)] dz, \quad (53)$$

which can be rewritten as

$$\bar{F}(m; \omega)|_{y=y_0} = \Delta^{(n)}(y_0, \omega) + \int g_{1d}[m^{(n)}; y_0, z, \omega] m^{(n)}(y_0, z) dz. \quad (54)$$

Substitution of equation (54) into equation (28) yields

$$\int g_{1d}[m^{(n)}; y_0, z, \omega] m^{(n+1)}(z) dz = e^{\text{obs}}(y_0, \omega) - e^{(n)}(y_0, \omega) + \int g_{1d}[m^{(n)}; y_0, z, \omega] m^{(n)}(z) dz. \quad (55)$$

The 2-D MT inverse problem is now solved by writing equation (55) as

$$\sum_i^{n_z} A_{ijk} m_{ij}^{(n+1)} = e_{jk}^{\text{obs}} - e_{jk}^{(n)} + \sum_i^{n_z} A_{ijk} m_{ij}^{(n)}, \quad j = 1, \dots, n_y, \quad k = 1, \dots, 2n_f, \quad (56)$$

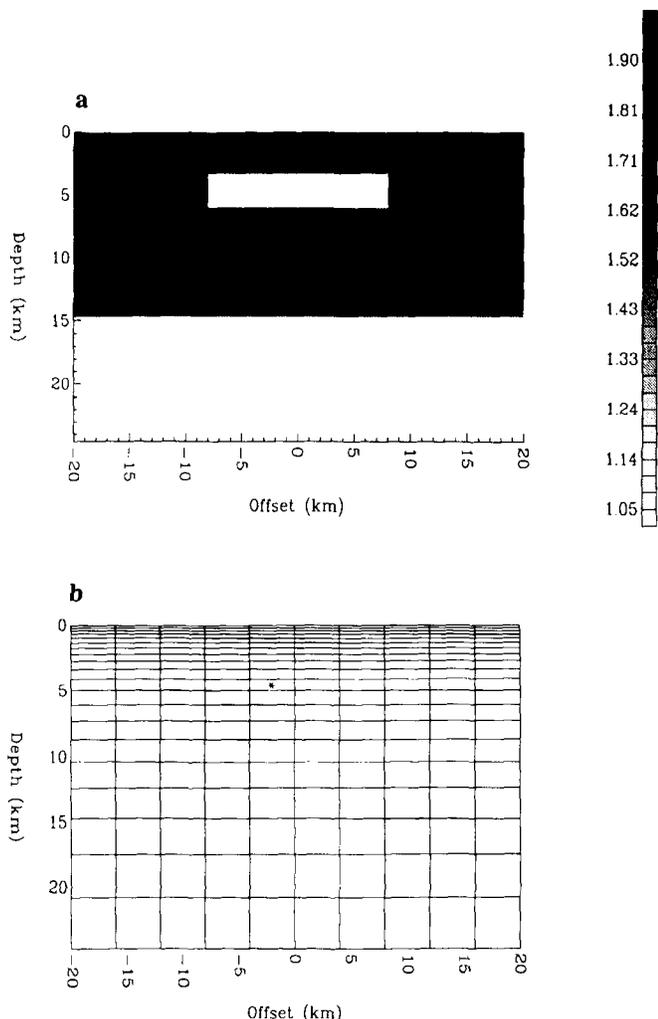
where  $A_{ijk}$  is the integral of the  $k$ th 1-D kernel function  $g_{1d}$  at model offset  $y_j$  over the  $i$ th depth partition. We minimize an objective function

$$\phi = \sum_k^{2n_f} \sum_j^{n_y} \left| \frac{e_{jk}^{\text{obs}} - e_{jk}^{(n)}}{\epsilon_{jk}} \right| + \beta \sum_i^{n_z-1} \sum_j^{n_y} |m_{i+1,j}^{(n+1)} - m_{i,j}^{(n+1)}| + \gamma \sum_i^{n_z} \sum_j^{n_y-1} w_i |m_{i,j+1}^{(n+1)} - m_{i,j}^{(n+1)}|, \quad (57)$$

where  $\epsilon_{jk}$  is the standard deviation of the datum  $e_{jk}^{\text{obs}}$ . The problem of minimizing (57) subject to the constraints in (56) is solved using linear programming (LP) techniques.

Having defined the inversion algorithms we now apply them to the inversion of 2-D MT data. A major goal in the presentation of these examples is to illustrate that 2-D conductivity structures can be generated using the AIM inversion even when only 1-D information is employed in the approximate inverse mapping. As such, we want to work with an algorithm which is most definitive in answering this question. This leads us to concentrate upon the AIM-DS formulation which minimizes of a global objective function of the conductivity. Examples of AIM inversions which make use of a base-model will be delayed until we treat the final example involving the recovery of two prisms buried in a two-layered earth.

To illustrate the AIM-DS approach in which a global objective function is minimized we begin with a simple 2-D model consisting of a 10  $\Omega$  m conducting prism in a resistive 100  $\Omega$  m background overlying a 10  $\Omega$  m conductive basement (Fig. 8a). The model is cellularized into a 20  $\times$  10 conductivity array as shown in Fig. 8(b). The forward modelled responses have been computed for the apparent resistivity and phase for the TE, TM polarizations and for the rotationally invariant determinant average response (Berdichevsky & Dmitriev 1976) which we denote by DET.



**Figure 8.** The single prism model consisting of a  $10 \Omega \text{ m}$  prismatic body buried in a  $100 \Omega \text{ m}$  background overlying a  $10 \Omega \text{ m}$  basement is shown in (a). The cellularization of the single prism model consisting of  $n_x \times n_z = 10 \times 20$  cells is given in (b). The cell marked by \* is the cell which has its resistivity decreased by 5 per cent to produce the sensitivity plot shown in Fig. (11).

These are shown in Fig. 9(a–c) for TE, TM, DET apparent resistivities and Fig. 10(a–c) for the phases. All computations use apparent resistivity and phase data from nine frequencies from 1.0 to 0.00033 Hz. For the purposes of inversion, the true data have an assigned standard deviation of 5 per cent for the apparent resistivities and  $2^\circ$  for the phases.

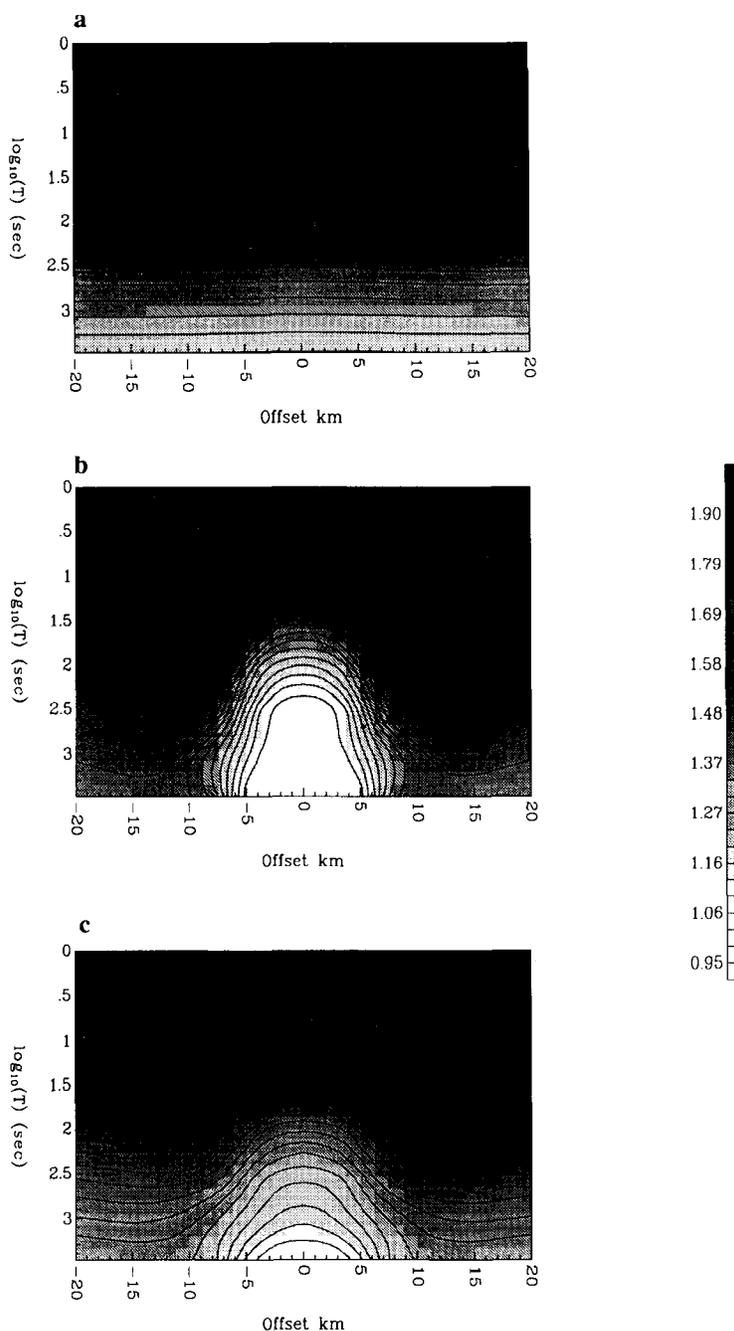
Before proceeding to the inversion of the data we motivate the assumption made above that the sensitivity function  $J(\sigma; y, y', z, \omega)$  is diagonally dominant with respect to  $y, y'$ . To do this we show, in Fig. 11 the difference between the DET response for the model of Fig. 8(a) and the same model with the cell marked by '\*' decreased in resistivity by 5 per cent. It can immediately be seen that the dominant change in response occurs above the perturbed cell.

Applying the AIM inversion as defined in equations (56, 57) to the TE mode data Figs 9(a) and 10(a), starting from a  $100 \Omega \text{ m}$  half-space, yields a series of models  $m^{(n)}$ . In

Fig. 12(a, b) we show  $m^{(n)}$  for  $n = 5, 10$ . Even at iteration  $n = 5$  the general structure of the true model is clearly visible: the conductive prism is resolved and the background and the basement are delineated. The primary discrepancy between the true model and the predicted model is that the predicted background resistivity is lower than the true background resistivity under the conductive prism. Comparison of Figs 12(a) and (b) show that the models  $m^{(5)}$  and  $m^{(10)}$  have very similar structure. This illustrates the typical behaviour of this AIM-DS inversion: the major features of the model are produced in the first few iterations and subsequent iterations produce smaller refinements. This is to be expected if the approximate inverse mapping  $\bar{F}^{-1}$  embodies the primary physics of the problem. In Fig. 12(c) (solid curve) we show the normalized  $\chi^2$  misfit, for the TE mode apparent resistivities, at each iteration. Specifically, at iteration  $n = 5$ ,  $\chi_N^2 = 0.21$ , and, at iteration  $n = 10$  the misfit  $\chi_N^2 = 0.09$ . Although this inversion used only the TE mode data it is of interest to display the TM mode misfit, i.e. the misfit between the true TM response and the TM response computed from models generated by this TE mode inversion. This is shown in Fig. 12(c) (dashed). The TM mode apparent resistivities show a much worse misfit than the TE mode apparent resistivities; this might be expected since they were not included in this inversion. Finally, it should be noted that this algorithm is extremely robust with respect to the starting model. For example, starting with a half-space model of  $10^4 \Omega \text{ m}$  or  $1 \Omega \text{ m}$  gives the same series of models for  $n \geq 5$ .

Next we apply the AIM inversion to the TM mode data (Figs 9b, 10b), starting from a  $100 \Omega \text{ m}$  half-space. This yields a series of models  $m^{(n)}$  of which we show  $m^{(n)}$  for  $n = 5, 10$  in Fig. 13(a, b). The general structure of the true model is clearly observed in the predicted model  $m^{(5)}$ . As the iterations proceed the amount of structure decreases (compare  $m^{(10)}$  with  $m^{(5)}$ ). This observation, together with the fact that the  $\chi^2$  misfit for models  $m^{(5)}$  and  $m^{(10)}$  are essentially the same, gives an important insight to this AIM inversion. This AIM-DS inversion, in the first few iterations, finds a model that greatly reduces the data misfit and in subsequent iterations the model is refined in accordance with the model norm in the objective function. In Fig. 13(c) (solid curve) we show the normalized  $\chi^2$  misfit for the TM mode apparent resistivities at each iteration. This is the analogous plot to Fig. 12(c). The TE mode misfit is shown in Fig. 13(c) (dashed). It is interesting that the TE mode data are fit somewhat better than the TM mode data even though only the latter were used in the inversion.

Finally, applying the AIM inversion to the DET mode data Fig. (9c, 10c), starting from a  $100 \Omega \text{ m}$  half-space, gives the models  $m^{(5)}$  and  $m^{(10)}$  shown in Fig. 14(a, b). Again we observe that the predicted model at iteration  $n = 5$  shows the general features of the true model. At iteration 10 the predicted model is an excellent 'minimum structure' representation of the true model. This emphasizes that the DET mode data is a good data mode for an AIM inversion which is based upon a composite of 1-D inversions; this supports the conclusions of Ranganayaki (1984) about the usefulness of DET data. In Fig. 14(c) (solid curve) we show the normalized  $\chi^2$  misfit, for the DET mode apparent resistivities, at each iteration. Specifically, at iteration  $n = 5$  the misfit  $\chi_N^2 = 0.51$  and at iteration  $n = 10$  the misfit



**Figure 9.** The apparent resistivity response for the single prism model as function of  $\log_{10}$  period and offset. The upper, middle, and lower plots show the TE, TM and DET mode responses respectively.

$\chi_N^2 = 0.26$ . Also shown in Fig. 14(c) (dashed) are the TE and TM mode misfits. These data fit nearly as well as they were in the individual inversion shown in Figs 12 and 13.

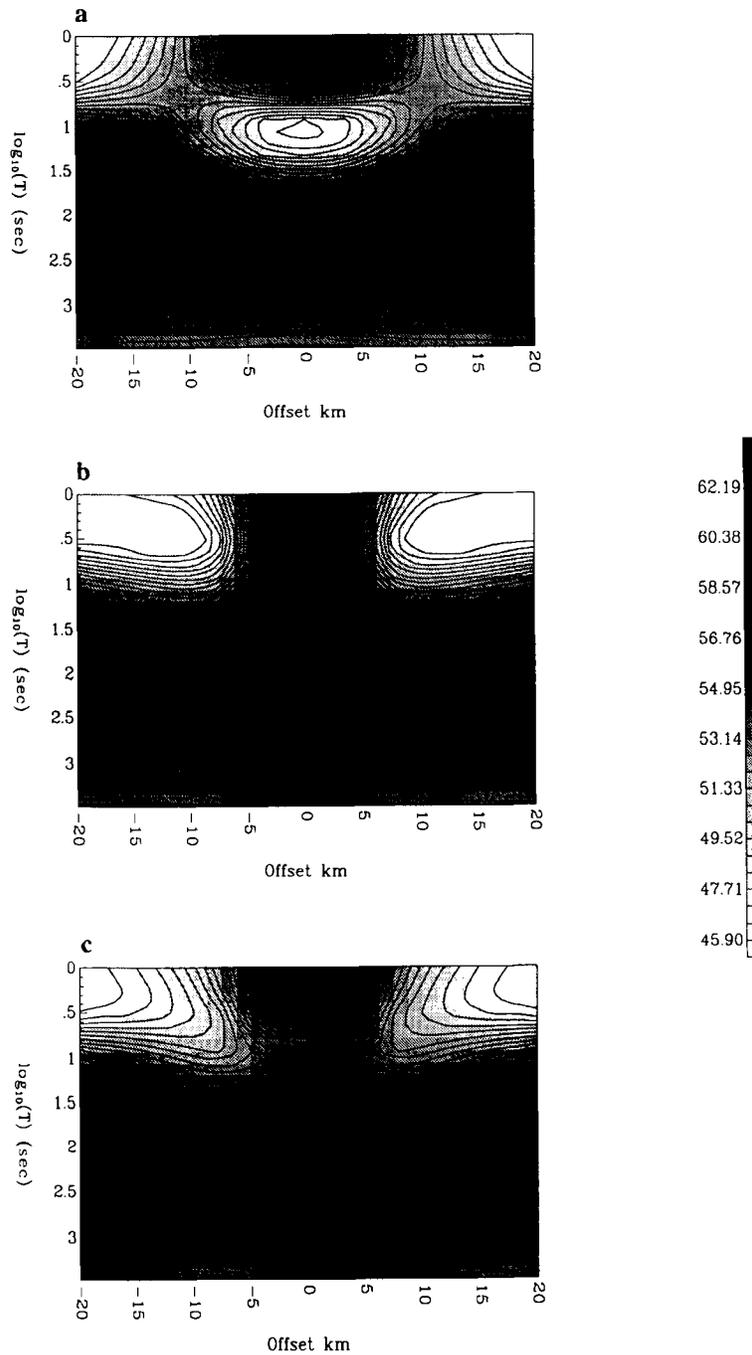
In order to demonstrate the robustness of this algorithm we have taken the DET mode data and added Gaussian independent noise of standard deviation 5 per cent to the apparent resistivities and  $2^\circ$  to the phases. We denote this as NDET data. Applying the AIM inversion to the NDET data starting from a  $100 \Omega \text{ m}$  half-space, gives the models shown in Fig. 15(a,b) for predicted models  $m^{(n)}$  at  $n = 5, 10$ . The misfit at iteration  $n = 5$  is  $\chi_N^2 = 1.8$  and at iteration  $n = 10$  is  $\chi_N^2 = 1.3$ . These results indicate that the algorithm is robust

and produces a good representation of the true model even in the presence of noise.

The preceding inversion results were all produced with the weights

$$w_i = \begin{cases} z_i \sqrt{z_{i+1} - z_i}, & i = 1, \dots, n_z - 1, \\ w_{n_z - 1}, & i = n_z. \end{cases} \quad (58)$$

There is an interesting theoretical question regarding the appropriate choice of weights and parameters for this problem; however, our approach on this point is pragmatic: we choose  $\beta$ ,  $\gamma$  by trial and error and choose  $w_i$  so that there



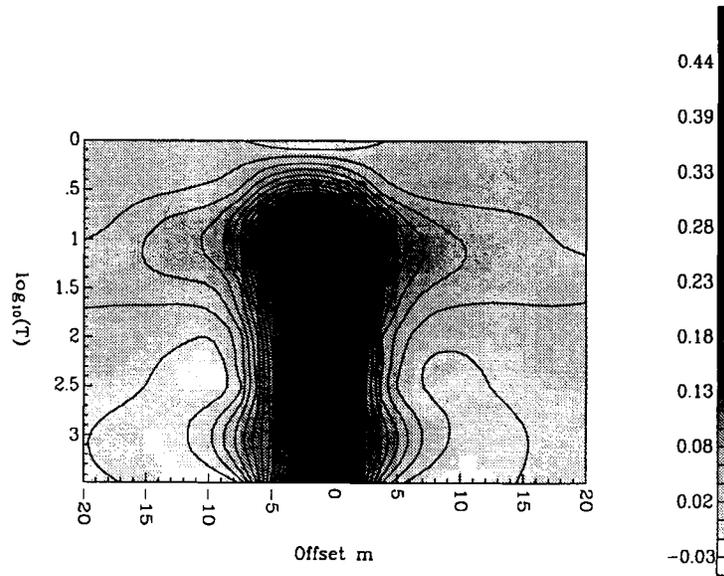
**Figure 10.** The phase response, in degrees, for the single prism model. The upper, middle, and lower plots show the TE, TM and DET mode responses respectively.

is less horizontal variation at greater depths and less variation between vertically thick cells than there is between thin cells. These choices lead to a preference for model variation near the surface. A useful consequence of this weighting is that static shift effects are automatically corrected in this inversion scheme because small near surface conductivity anomalies are favoured relative to large deep anomalies. This has proved extremely useful in the inversion of field data with this algorithm.

To summarize our observations from the single prism model we note first that comparison of the inversion of TE,

TM, DET mode data clearly indicate that the DET mode data is an excellent choice of data for AIM inversion which uses an approximate inverse mapping based on a composite of 1-D inversions. Second, we note that the AIM algorithm defined by equations (56), (57) shows rapid convergence to a model that 'fits the data' then slower convergence to a model with 'desired structure'. Third the algorithm is robust in the presence of noise.

The preceding inversions demonstrated how the AIM technique could be formulated to invert 2-D magnetotelluric data. The single prism test model was useful in illustrating



**Figure 11.** The relative change in apparent resistivity when the cell marked in Fig. (8b) is decreased in resistivity by 5 per cent. The quantity  $[\log_{10} \rho_A(\text{pert}) - \log_{10} \rho_A(\text{base})] / \log_{10} \rho_A(\text{pert})$  is plotted where *pert* refers to the perturbed model and *base* refers to the initial model.

several important points about the inversion algorithm and about 2-D magnetotellurics in general; however, it was limited to a relatively small scale having only 200 parameters and 180 data. We remove this final limitation and now apply the AIM inversion to a larger scale problem having 900 data and 1500 model parameters. We use the same inversion scheme as in the single prism model; however, in this example the model we choose has  $n_y = 50$  horizontal cells and  $n_z = 30$  vertical cells. The model consists of a 1000  $\Omega$  m resistive prism and a 10  $\Omega$  m conductive prism in a 100  $\Omega$  m background all overlying a 10  $\Omega$  m conductive basement (Fig 16a). This model was chosen mainly because of its similarity with models already used by other authors (Smith & Booker 1988b; deGroot-Hedlin & Constable 1989). It is hoped that a comparison of the same model inverted by different techniques will provide information about the strengths and weaknesses of different methods. The cellularization of the model (Fig. 16b) is similar to that used in the single prism model: logarithmically increasing cell thickness with depth; constant cell width in the region of interest; a 'hidden' border of cells of exponentially increasing dimension to provide appropriate boundary conditions. When referring to data associated with this model we consider nine frequencies in the range 1.0 to 0.00033 Hz.

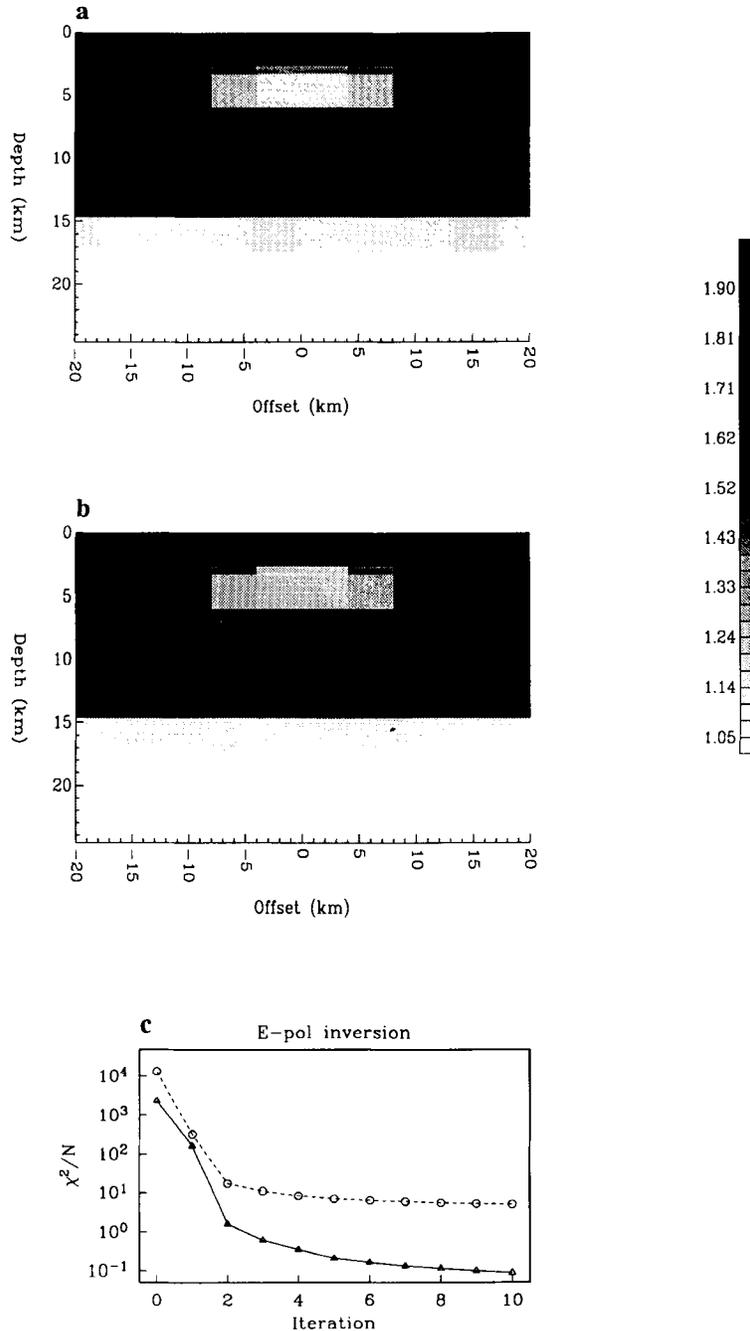
The single prism test model indicated that the DET mode data was a good choice of data for inversion using an approximate inverse mapping based on a composite of 1-D inversion. Consequently we invert only the determinant average mode data for the two prism model. The data to be inverted,  $e^{\text{obs}}$ , are shown in Fig. 17(a,b) for the apparent resistivity and phase. For this model we invert two sets of data; the exact DET data, and data with Gaussian independent noise added, NDET data. The noise is the same as the previous examples: the standard deviations are 5 per cent on the apparent resistivities;  $2^\circ$  on the phases.

We begin by applying the AIM-MS and AIM-DS inversion procedures when the approximate inverse mapping incorporates a base-model. To estimate a

base-model we carried out a 1-D AIM-DS inversion at each site (using the algorithm that generated the results in Fig. 7) and averaged the results to find a single conductivity depth function. This 1-D conductivity was designated as the base-model for each station. For this example, the resultant base-model was nearly identical to the 1-D conductivity represented by the surface layer and underlying half-space. This similarity resulted because the effects of the conductive and resistive prisms tended to cancel in the upper layer, and because the LP minimum variation model recovered a layer over a half-space when data away from the prisms were inverted.

The accurate DET data were inverted with the AIM-MS algorithm described by equation (15) where  $\bar{F}^{-1}$  is the composite of 1-D inversions obtained by minimizing equation (51). The model obtained after six iterations is shown in Fig. 18(a). The normalized  $\chi^2$  misfit, again assuming an error of 5 per cent on the apparent resistivities and  $2^\circ$  on the phases, is plotted in Fig. 18(c). By iteration 6  $\chi_N^2 = 0.15$  for the apparent resistivities and  $\chi_N^2 = 0.096$  for the phases. Overall, the correspondence between Figs 18(a) and 16(a) is quite good. Both prisms are well imaged, and the background conductivities in the upper layer and in the bottom half-space are well approximated. There are artifacts however. The conductive prism is smeared out in both the lateral and vertical directions and the resistive prism has moved somewhat toward the surface and exhibits a drop-out beneath.

The application of the AIM-DS algorithm which uses equation (27), specifies  $\bar{F}^{-1}$  to be the composite of 1-D inversions obtained by minimizing equation (51), and uses the same base-model as described above, is shown in Fig. 18(b). The result is somewhat better than that observed in Fig. 18(a). The prisms are exceedingly well represented. The normalized  $\chi^2$  misfit error for successive iterations is shown in Fig. 18(d). The convergence is monotonic and by iteration 6,  $\chi_N^2 = 0.098$  for the apparent resistivities and  $\chi_N^2 = 0.031$  for the phases.

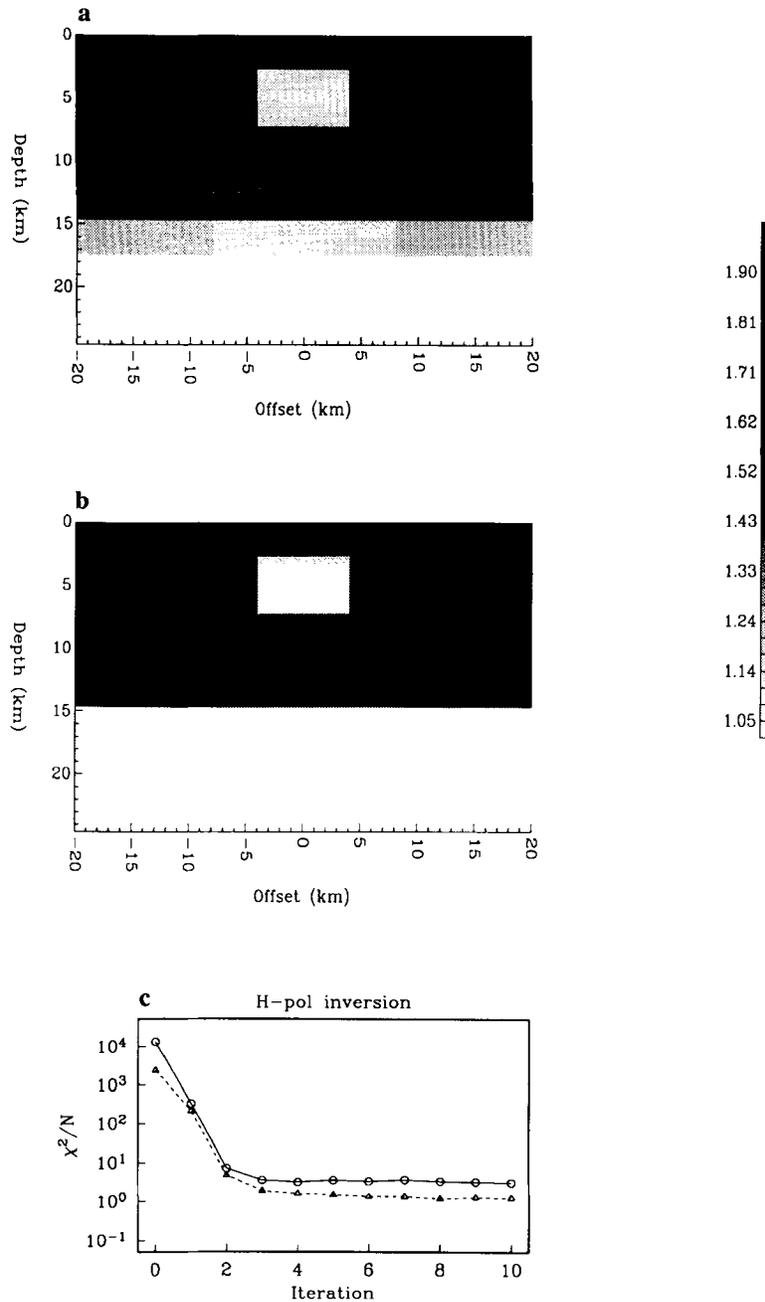


**Figure 12.** The TE mode inversion for the single prism model. The upper plot shows the predicted model at iteration 5 and the middle plot shows the predicted model at iteration 10. The lower plot shows the normalized  $\chi^2$  misfit as a function of iteration. The dashed line indicates the fit to the TM mode data even though they are not included in the inversion.

To summarize, in these inversions where the base-model is essentially the background conductivity, the AIM technique has worked well when the approximate inverse mapping is relegated to be composed of 1-D inversions. Another attractive feature of the two AIM inversions portrayed in Fig. 18 is the computational efficiency. Application of an approximate inverse mapping (which consisted of 50 1-D inversions) required only about 500 s on a SUN 4/370 workstation. A complete iteration in the inversion, including accurate forward modelling, required about 1000 s on average.

We now turn our attention to the AIM-DS inversion employing the approximate forward mapping of equation (54). In contrast to the preceding inversion which was based on optimizing the objective function of equation (51), this inversion is based on optimizing the objective function of equation (57) and *does not* require a base model.

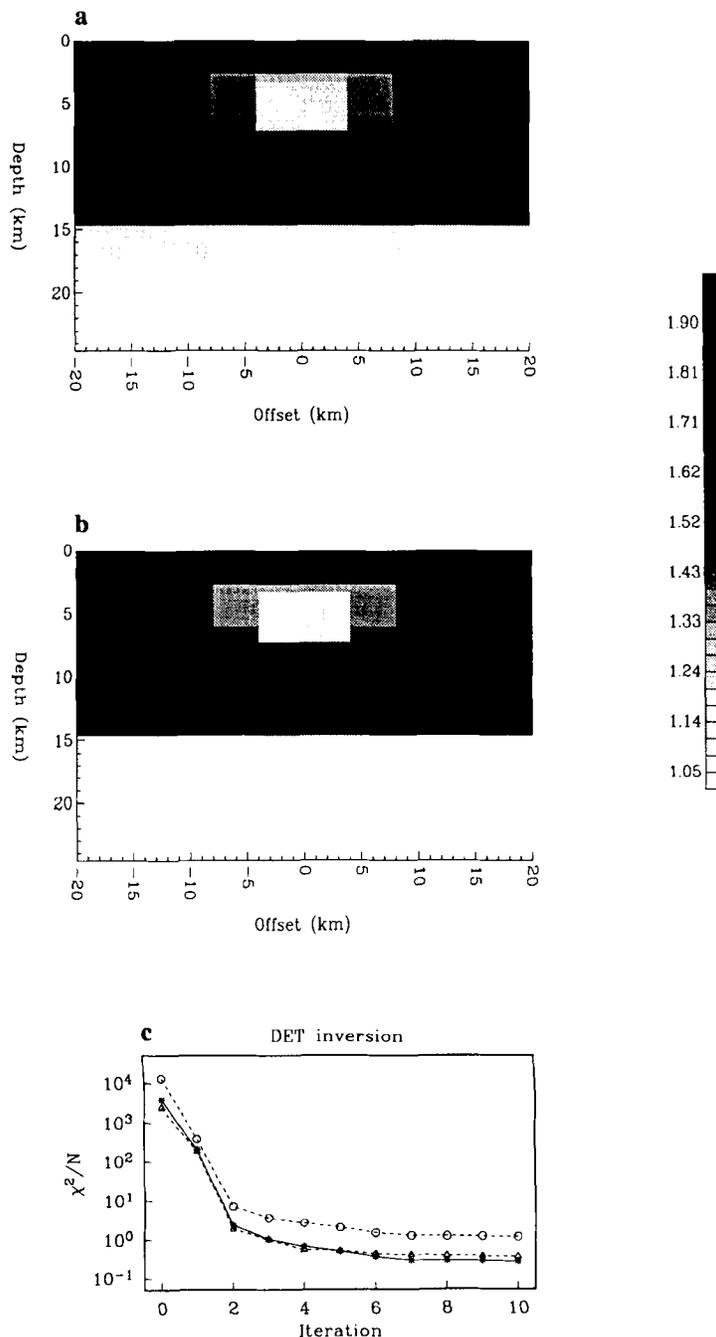
First, we apply the inversion algorithm (equations 54, 57) to the exact data, starting from a uniform half-space. The models,  $m^{(n)}$ , produced by the inversion for iterations  $n = 5, 10$  are shown in Fig. 19(a, b). The inversion was performed with the weights  $w_i$  as defined in equation (58).



**Figure 13.** The TM mode inversion for the single prism model. The upper plot shows the predicted model at iteration 5, the middle plot shows the predicted model at iteration 10. In the lower plot the normalized  $\chi^2$  misfit as a function of iteration is indicated by the solid line. The dashed line indicates the fit to the TE mode data even though they are not included in the inversion.

After five iterations the predicted model clearly shows the structure of the true model. The conductive and resistive prisms are resolved and the basement is clearly delineated. The normalized misfit at this iteration is  $\chi_N^2 = 0.12$  for the apparent resistivity, and  $\chi_N^2 = 0.04$  for the phases. Clearly this is a very good fit to the data. Fig. 19(c) shows the normalized misfit as a function of iteration. After further iterations, the inversion produces models with less structure. For example, the model produced at iteration  $n = 100$  more clearly shows the two prisms and the conducting basement, and has very little extraneous structure. The normalized misfit for the apparent resistivity is  $\chi_N^2 = 0.012$  and for the

phase  $\chi_N^2 = 0.0008$ . The final inversion result, Fig. 19(b), 'fits the data' and shows an excellent resemblance to the true model Fig. 16(a). The background and basement resistivities are uniform and well estimated; however, there are some discrepancies in the shape and resistivities of the two prisms. First, the predicted conductive prism has slightly less depth extent and slightly higher conductivity than the true conductive prism. This is due to insufficient frequency domain sampling and the tendency of the norm equation (57) to minimize structure. Second, for similar reasons, the predicted resistive prism has about twice the depth and lateral extent of the true resistive prism and about 1/4 the

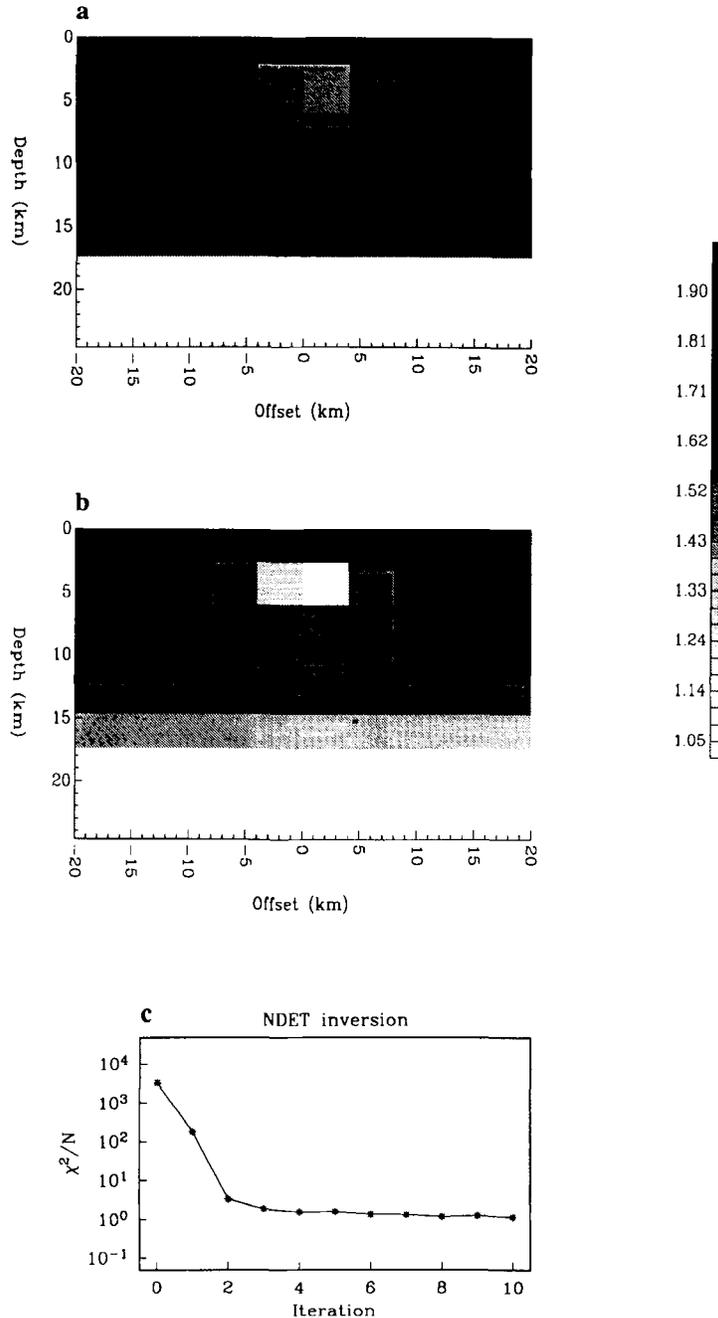


**Figure 14.** The DET mode inversion for the single prism model for data without noise. The upper plot shows the predicted model at iteration 5, the middle plot shows the predicted model at iteration 10. The lower plot shows the normalized  $\chi^2$  misfit as a function of iteration. The solid line refers to the DET data; the dashed lines monitor how well individual TE and TM data were fit. The  $\Delta$ 's and  $\circ$ 's correspond to the TE and TM modes respectively.

conductivity. It is also well known that a resistive inclusion is hard to image in the TE mode and consequently, when combined with the TM mode data to produce the DET mode data, is somewhat blurred.

Second, we apply the inversion algorithm (equations 54, 57) to the NDET data, i.e. the data with noise added. The results of inverting the NDET data are shown in Fig. 20. The models,  $m^{(n)}$  are shown for iterations  $n = 11, 37$ . After 11 iterations with  $w_i$  as defined by equation (58), we obtain the model shown Fig. 20(a). It generates a

normalized misfit of  $\chi_N^2 = 1.2$  for the apparent resistivities, and of  $\chi_N^2 = 1.0$  for the phases. Already the general structure of the true model Fig. 16(a) is clearly evident in Fig. 20(a). After 37 iterations a model is obtained with less structure and better resolution of the two prisms. The effect of the noise on the inversion has been to generate small perturbations in the background resistivity near the surface, while leaving the large-scale structure practically unchanged. The model  $m^{(37)}$  has a misfit  $\chi_N^2 = 1.2$  for the apparent resistivities and  $\chi_N^2 = 1.0$  for phases. In Fig. 20(c) we show

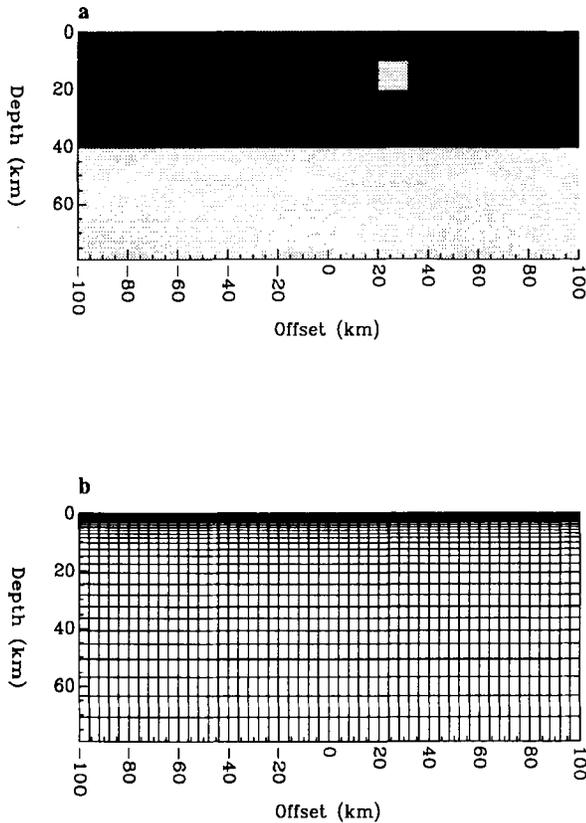


**Figure 15.** The NDET mode inversion for the single prism model for data with noise. The upper plot shows the predicted model at iteration 5, the middle plot shows the predicted model at iteration 10. The lower plot shows the normalized  $\chi^2$  misfit of the apparent resistivities as a function of iteration.

the normalized  $\chi^2$  misfit for the apparent resistivity and phase as a function of iteration. This inversion result shows that even in the presence of noise the characteristic behaviour of this AIM algorithm is the same; the first few iterations use the primary physics of the approximate inverse mapping  $\bar{F}^{-1}$  to produce a model which approximately fits the data, then subsequent iterations produce a model with less structure.

An analysis of the inversion results for the two prism model shows that this AIM-DS inversion, which minimizes a

2-D objective function, can be used very successfully on large-scale problems. We emphasize that only 1-D kernel functions have been used thereby eliminating the need for the computationally intensive process of computing 2-D sensitivities. The two prism model inversion with 1500 parameters and 900 data leads to an LP solution which has to be computed at each iteration and this is computationally demanding, taking on average approximately 1 hr of SUN4/370 CPU time per iteration. Even though reusing the LP basis obtained from the previous iteration decreases



**Figure 16.** The two prism model consisting of a  $1000 \Omega \text{ m}$  prismatic body (left) and a  $10 \Omega \text{ m}$  prismatic body (right) buried in a  $100 \Omega \text{ m}$  background overlying a  $10 \Omega \text{ m}$  basement is shown in (a). The cellularization of the single prism model consisting of a  $n_z \times n_x = 30 \times 50$  cells is given in (b).

dramatically the CPU time per iteration as the inversion progresses, work is still needed to optimize this phase of the inversion.

## 6 DISCUSSION AND CONCLUSIONS

The AIM inverse offers considerable promise as a general method for solving large-scale inverse problems. The positive aspect is that which motivated the research to begin with, viz. the development of a computationally efficient inversion method which requires neither multidimensional linearization of the equations nor subsequent solution of a very large system of matrix equations. The technique is still in its infancy; however, from the preliminary work done to date there are some general conclusions that can be drawn, and strengths and weaknesses that can be delineated.

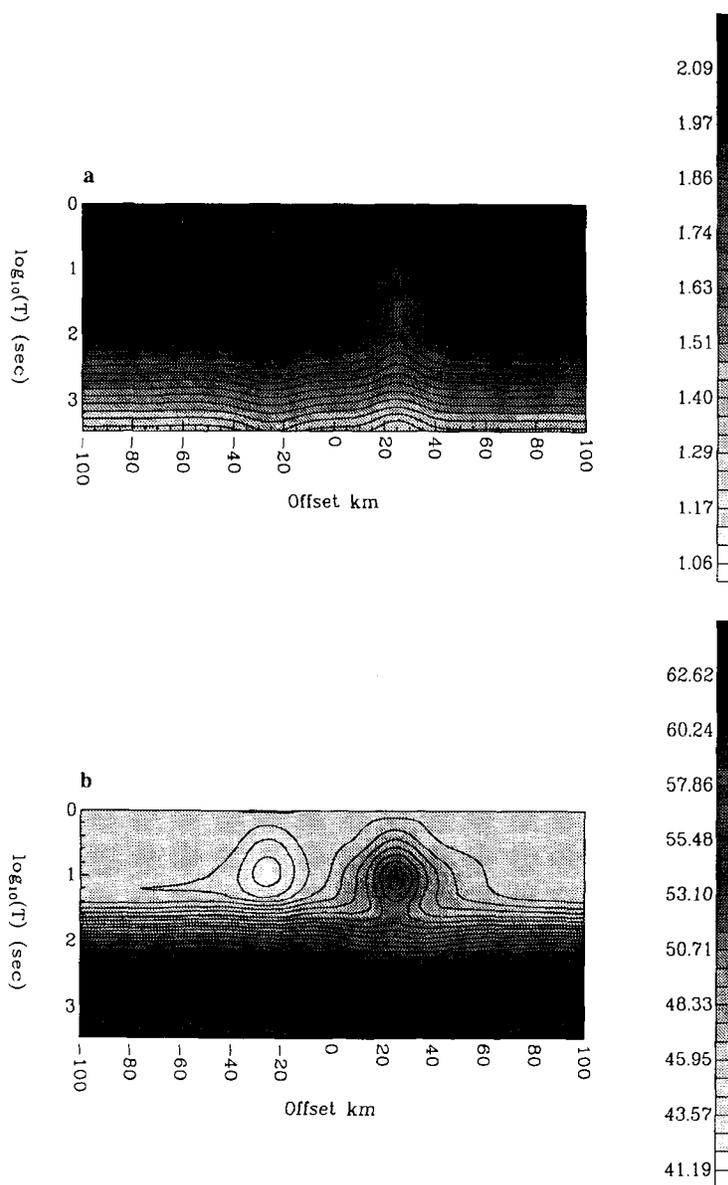
In the general formulation of the AIM inversion method we began with the concept of an approximate inverse mapping. This approximate mapping was used in two general iterative methods: AIM-MS where iterative corrections were made in model space and AIM-DS where iterative corrections were made in data space. Although both methods can produce good results we prefer AIM-DS. There are at least two negative aspects of AIM-MS that lead us to this conclusion. Firstly, there is no direct way to incorporate a desired norm of the final model into the AIM-MS inversion. Perturbation functions are successively

added to generate a final solution; this tends to roughen the model and is undesirable. In the AIM-DS formulation a model which minimizes a global objective function can be found; this is a major advantage. The second negative aspect of AIM-MS is that there are computational difficulties because the model perturbation is the difference between the application of a specific  $\tilde{F}^{-1}$  to the observations and to the predicted data, viz. equation (11). The definition of  $\tilde{F}^{-1}$  must include any regularization that is applied, and any model produced by application of  $\tilde{F}^{-1}$  inherits artifacts of that regularization. It is computationally difficult to ensure that the regularization is compatible for the two applications of  $\tilde{F}^{-1}$  and that an appropriate perturbation step is thereby computed.

There are many questions about existence and uniqueness of a solution and convergence that need to be explored. At a basic level however, it is clear that existence, uniqueness, and convergence to a solution using an approximate inverse mapping,  $\tilde{F}^{-1}$ , are related to the fixed points of the model space and data space operators  $\tilde{m}^{\text{obs}} + \Delta_m$  and  $e^{\text{obs}} + \Delta_e$  and the properties of these operators in the neighbourhood of fixed points. Whilst an understanding of these properties is ultimately desirable, the complexity of  $\tilde{F}^{-1}$  is usually such that an empirical approach to these questions, based on physical understanding of the process  $F$ , is all that is possible for realistic problems. We have shown in this paper that such an approach is feasible.

The formulation of the AIM-DS method and the realization that convergence is ultimately tied to properties of fixed point algorithms have provided essential insight into model construction in non-linear inverse problems. In an attempt to extremize a global objective function of the conductivity, Oldenburg (1983) linearized the 1-D equations and rearranged them so that modified data were a linear functional of the model to be determined. He gave no justification for why, and under what conditions, that reformulation might converge. The derivation of the AIM-DS algorithm presented here offers the lacking insight. Oldenburg's approach is properly thought of as being a special case of the AIM-DS algorithm where the approximate forward mapping is given by the sum of a constant and an inner product of the Fréchet kernel with the model. Convergence of the solution will depend upon the fixed point properties of  $e^{\text{obs}} + \Delta_e$ . Consequently, despite the fact that global optimization seems to have worked well for the 1-D MT problem when the Fréchet kernel is used in a forward mapping, the intriguing question remains as to whether different approximate forward mappings might be even better.

The general question 'what is the best approximate inverse mapping?' for any particular problem is unanswerable at this time. We have adopted a rather philosophical approach and attempted to break the problem into 'primary' and 'residual' physics. Our examples with electromagnetic induction have illustrated that this may be a good choice and that it provides a solid foundation for the development. This may mean that the complete inversion of an multidimensional problem can be achieved using, as an approximate inverse mapping, a full inversion in a lower dimension. The results achieved in the 2-D MT inversion are rather remarkable in this regard. The approximate inverse mappings used only 1-D information. No explicit

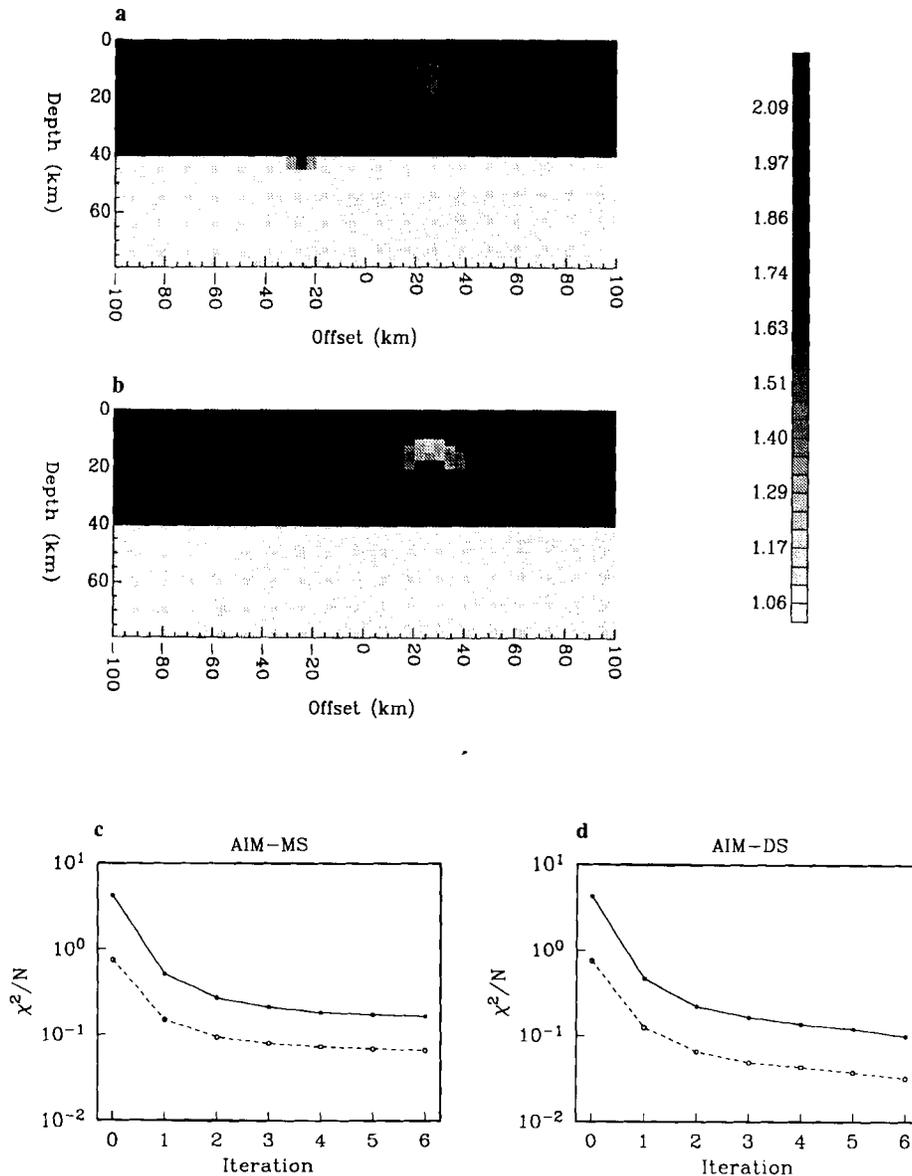


**Figure 17.** The DET mode apparent resistivity (upper) and phase (lower) response for the two prism model as function of  $\log_{10}$  period and offset.

information about lateral conductivity variation was included; nor was there any incorporation of 2-D physical effects such as the accumulation of electrical charges and the generation of laterally varying currents which are present in TM and TE mode induction problems. Yet the algorithms converged to excellent 2-D models. With respect to the choice of an approximate inverse mapping used in the AIM-DS algorithm incorporating a linear approximate forward mapping, it would seem that even simple modifications to the 1-D inversion, such as incorporating the 2-D electric fields (which are available at no cost from the forward modelling) into the kernel functions (e.g. Smith & Booker 1988b) would improve the mapping; including an approximation to the true 2-D sensitivities should do even better. We have deliberately chosen not to investigate such refinements here so that the capabilities of the AIM

inversion could be explored when the simplest approximate inverses were used.

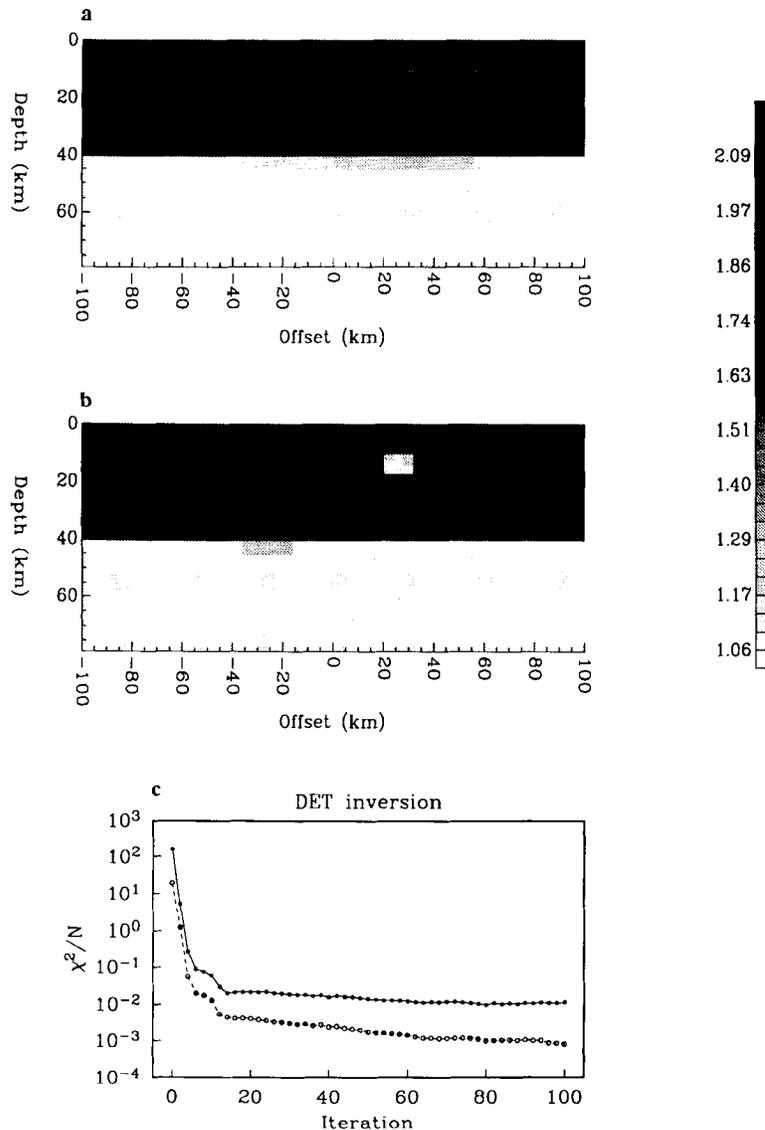
The formulations of the inversion which require a base-model have both advantages and disadvantages. The disadvantage becomes immediately apparent in those circumstances where nothing is known *a priori* about the model. In such cases it is difficult to specify the base-model. Since the constructed model will be influenced by the base-model, a poor estimate will manifest itself in a poor inversion result. The advantages of the base-model inclusion however, are that the inversion can be carried out very quickly, as illustrated by the 2-D inversions presented here. Also, there are circumstances, such as in near-surface investigations where borehole information is available, when sufficient information may be available to construct a base-model.



**Figure 18.** The AIM-MS and AIM-DS base-model inversions for the two prism model. The inversion makes use of a 1-D base-model obtained by averaging minimum structure 1-D inversions of the data. The AIM-MS model shown in (a) was obtained after six iterations. The AIM-DS model, obtained after six iterations, is shown in (b). The normalized  $\chi^2$  misfit for the apparent resistivities (solid line) and phases (dashed line) as a function of iteration for the AIM-MS and AIM-DS inversions are shown in (c) and (d) respectively.

In our opinion there is a hierarchy of techniques to be used in solving any inverse problem. The best involves generation of the complete sensitivity matrix, using these kernels as an approximate forward mapping in an AIM-DS algorithm, and computing a new model by minimizing an objective function subject to the data constraints. This involves solving a large system of equations. This approach, which is the straightforward extension of techniques used to solve 1-D inverse problems is conceptually the simplest to implement but is computationally demanding. As the problem becomes larger it becomes necessary to make sacrifices. Depending upon the problem, it may be that computation of the sensitivities is too demanding to be carried out, or it may be the inversion of the large matrix

which is the limiting factor. We have illustrated in the 2-D MT inversion an example where the computations of the complete 2-D sensitivities may be omitted and that an objective function minimization leads to an excellent model. That is, the model norm appears to be able to drive the algorithm to the right kind of model even if very approximate sensitivities are supplied. At an even lower level of computational requirements, we have inverted 2-D MT data by carrying out many individual 1-D inversions when a base-model was incorporated into the objective function. In this approach neither correct sensitivities nor the solution of a large matrix system was required. For the 2-D MT problem where 1500 parameters were estimated from 900 data, we found all of the inverse solutions



**Figure 19.** The DET mode inversion for the two prism model for data without noise. The upper plot shows the predicted model at iteration 5, the middle plot shows the predicted model at iteration 97, and the lower plot shows the normalized  $\chi^2$  misfit as a function of iteration. The solid and dashed lines respectively refer to the misfit of apparent resistivities and of phases.

presented here were completely tractable using a deskside workstation.

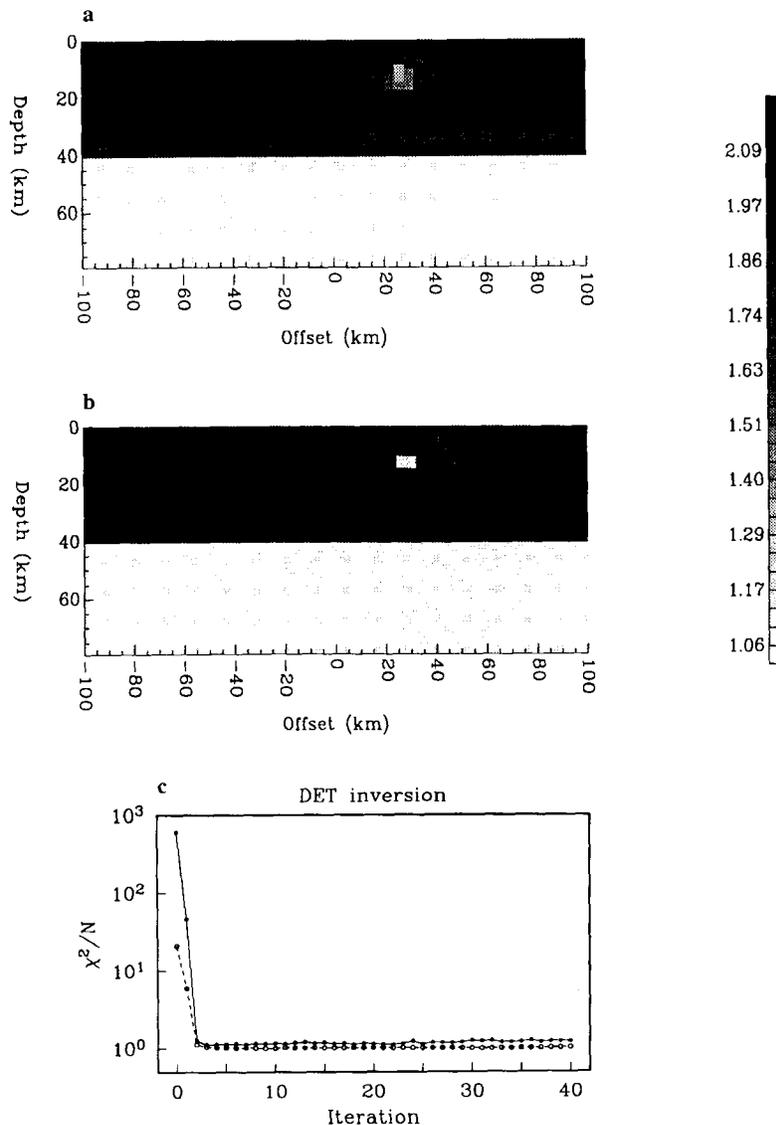
Although this paper has dealt with the AIM inversion applied to the magnetotelluric problem we emphasize that this inversion technique can be applied to any inverse problem provided that two crucial components are available. The first component is the ability to apply  $F$  and thereby to compute accurate forward model responses. The second necessary component is the existence of an approximate inverse mapping,  $\tilde{F}^{-1}$ . Thus any inverse problem for which  $F$  is known and for which a reasonable guess can be made for  $\tilde{F}^{-1}$  is amenable to the AIM inversion technique.

In summary, the results of the AIM inversion are very encouraging and lead us to believe that large-scale geophysical inverse problems can be solved without the

need to carry out a complete linearization and without the need to solve a large system of equations. We are therefore optimistic that our inversion procedures ultimately will prove useful in solving inverse problems which involve thousands of data and models with millions of unknowns.

#### ACKNOWLEDGMENTS

The authors are grateful to S. Dosso for many valuable conversations concerning 1-D magnetotelluric inversion. The authors also thank K. Whittall, Y. Li, D. Aldridge and the reviewers for suggestions which improved the clarity of the manuscript. This work was supported by an NSERC grant 5-84270.



**Figure 20.** The NDET mode inversion for the two prism model for data with noise. The upper plot shows the predicted model at iteration 11, the middle plot shows the predicted model at iteration 37. The normalized  $\chi^2$  misfit as a function of iteration for apparent resistivities (solid line) and phases (dashed line) are shown in the bottom panel.

## REFERENCES

- Berdichevsky, M. N. & Dmitriev, V. I., 1976. Basic principles of interpretation of magnetotelluric sounding curves, in *Geoelectric and Geothermic Studies, KAPG Geophysical Monograph*, ed. Adam, A., Akademiai Kiado, Budapest.
- Bostick, F. X. Jr, 1977. A simple almost exact method of magnetotelluric analysis: Workshop on Electrical Methods in Geothermal Exploration, U.S. Geol. Surv. Contract No. 14080001-G-359.
- 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.
- Defrise, M. & De Mol, C., 1987. A note on the stopping rules for iterative regularization methods and filtered SVD, in *Advances in Electronics and Electron Physics, Supplement 19*, ed. Sabatier, P. C., Academic Press.
- deGroot-Hedlin, C. & Constable, S. C., 1990. Occam's Inversion to generate smooth, two-dimensional models from magnetotelluric data, *Geophysics*, **55**, 1613–1624.
- Dosso, S. E. & Oldenburg, D. W., 1989. Linear and non-linear appraisal using extremal models of bounded variation, *Geophys. J. Int.*, **99**, 483–495.
- Gamble, T. D., 1983. Simple one-dimensional magnetotelluric inversions, in *Expanded Abstracts, 53rd Annual SEG meeting*, Las Vegas, USA, pp. 148–151, SEG, Tulsa, OK.
- Goldberg, S., Loewenthal, D. & Rotstein, Y., 1982. An improved algorithm for magnetotelluric and direct current data interpretation, *J. Geophys.*, **50**, 151–158.
- Hagedoorn, J. G., 1959. The plus-minus method of interpreting seismic refraction sections, *Geophys. Prosp.*, **7**, 158–182.
- Jupp, D. L. B. & Vozoff, K., 1975. Stable iterative methods for the inversion of geophysical data, *Geophys. J. R. astr. Soc.*, **42**, 957–976.
- Kaufman, A. A. & Keller, G. V., 1983. *Frequency and Transient Soundings*, Elsevier, Amsterdam.
- Levenberg, K., 1944. A method for the solution of certain

- non-linear problems in least squares, *Quart. Appl. Math.*, **2**, 164–168.
- Madden, T. R., 1972. Transmission systems and network analogies to geophysical forward and inverse problems, *Tech. Rep. No. 72-3*, Department of Earth and Planetary Sciences, MIT.
- Marquardt, D. W., 1970. Generalized inverses, ridge regression, biased linear estimation, and non-linear estimation, *Technometrics*, **12**, 591–612.
- Marsten, R. E., 1981. The design of the XMP linear programming library, *ACM Trans. Math. Software*, **7**, 481–497.
- Niblett, E. R. & Sayn-Wittgenstein, C., 1960. Variation of electrical conductivity with depth by the magneto-telluric method, *Geophysics*, **25**, 998–1008.
- 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.
- Parker, R. L., 1980. The inverse problem of electromagnetic induction: existence and construction of solutions based on incomplete data, *J. geophys. Res.*, **85**, 4421–4428.
- Ranganayaki, R. P., 1984. An interpretive analysis of magneto-telluric data, *Geophysics*, **49**, 1730–1748.
- Ryu, J., Morrison, H. F. & Ward, S. H., 1972. Electromagnetic depth sounding experiment across Santa Clara valley, *Geophysics*, **37**, 351–374.
- Smith, J. T., 1988. Rapid inversion of multi-dimensional magnetotelluric data, *PhD thesis*, Geophysics Program, University of Washington, Seattle.
- Smith, J. T. & Booker, J. R., 1988a. Magnetotelluric inversion for minimum structure, *Geophysics*, **53**, 1565–1576.
- Smith, J. T. & Booker, J. R., 1988b. Two and three dimensional inversion of magnetotelluric data, presented as the *Electromagnetic Induction Workshop*, June 1988, Ensenada, Mexico.
- Stratton, J. A., 1941. *Electromagnetic Theory*, McGraw-Hill, New York.
- Tikhonov, A. N., & Arsenin, V. Y., 1977. *Solutions of Ill-posed Problems*, ed. Fritz, J., John Wiley & Sons, New York.
- Vozoff, K., 1972. The magnetotelluric method in the exploration of sedimentary basins, *Geophysics*, **37**, 98–141.
- Ward, S. H. & Hohmann, G. W., 1988. Electromagnetic theory for geophysical applications, in *Electromagnetic Methods in Applied Geophysics Vol 1, Theory*, ed. Nabighian, M. N., SEG publication, Tulsa OK.
- 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, Tulsa, OK, in press.
- Word, D. R., Smith, H. W. & Bostick, F. X. Jr, 1970. An investigation of the magnetotelluric impedance method, *Electr. Geophys. Res. Lab. Tech. Rep. No. 82*, University of Texas, Austin.
- Zohdy, A. A. R., 1974. Use of Dar Zarrouk curves in the interpretation of vertical electrical sounding data, *US Geol. Surv. Bull.*, **1313-D**, 1–41.