GEOPHYSICS, VOL. 55, NO. 12 (DECEMBER 1990); P. 1613–1624, 10 FIGS., 1 TABLE.

Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data

C. deGroot-Hedlin* and S. Constable*

ABSTRACT

Magnetotelluric (MT) data are inverted for smooth 2-D models using an extension of the existing 1-D algorithm, Occam's inversion. Since an MT data set consists of a finite number of imprecise data, an infinity of solutions to the inverse problem exists. Fitting field or synthetic electromagnetic data as closely as possible results in theoretical models with a maximum amount of roughness, or structure. However, by relaxing the misfit criterion only a small amount, models which are maximally smooth may be generated. Smooth models are less likely to result in overinterpretation of the data and reflect the true resolving power of the MT method. The models are composed of a large number of rectangular prisms, each having a constant conductivity. A priori information, in the form of boundary locations only or both boundary locations and conductivity, may be included, providing a powerful tool for improving the resolving power of the data. Joint inversion of TE and TM synthetic data generated from known models allows comparison of smooth models with the true structure. In most cases, smoothed versions of the true structure may be recovered in 12–16 iterations. However, resistive features with a size comparable to depth of burial are poorly resolved. Real MT data present problems of non-Gaussian data errors, the breakdown of the two-dimensionality assumption and the large number of data in broadband soundings; nevertheless, real data can be inverted using the algorithm.

INTRODUCTION

It is well known that magnetotelluric (MT) inversion with a finite data set is nonunique and therefore an infinite number of conductivity structures exist which fit the data, if any exist at all. Despite this, a common approach to fitting a two-dimensional (2-D) MT data set is to construct a cross-section of the area based on prior geological knowledge and then to solve for the conductivities by least-squares inversion (Jupp and Vozoff, 1977) or by a trial-and-error forward modelling method (EMSLAB group, 1988). These solutions are highly dependent upon the model parameterization and the prior assumptions about the geology. Finding a model based on an assumed geological structure which has an adequate fit to the data may tempt one to believe that features appearing in the model are necessary rather than merely consistent with the data and one might argue that in using this method little is learned independently from the MT data about structure in the Earth.

A more objective approach is to solve for model structure by overparameterizing the model; that is, dividing the model into more blocks than there are degrees of freedom in the data, and then solving for the conductivities. However, while underparameterizing the model may suppress significant structure, overparameterizing the model and conducting a least-squares inversion will lead to unstable solutions which contain large oscillations. In the extreme case, the least-squares best-fitting model possesses structure which is rougher than is physically possible (e.g., Parker, 1980). Imposing a smoothness constraint on the model may stabilize the solution (Rodi et al., 1984; Sasaki, 1989), but unless the smoothest model is explicitly sought, features may still appear which are not required by the data.

The approach presented in this paper is to find models fitting the data which are extreme in the sense of having the

Manuscript received by the Editor August 4, 1989; revised manuscript received May 23, 1990.
*Institute of Geophysics and Planetary Physics, Scripps Institution of Oceanography, A-025, La Jolla, CA 92093.

minimum possible structure. A tenet of modern science known as Occam's razor is followed: a simple solution is preferable to one which is unnecessarily complicated. The approach closely follows that of Constable et al. (1987) for 1-D EM inversion in that the smoothest possible model is sought at a given level of misfit for an overparameterized model. Although these models are not necessarily closer to the truth than any other models which fit the data, they give lower bounds on the amount of structure required. It is then likely that the true Earth is at least as rough as the models. Also, smooth models give an indication of the resolving power of the EM method, since the data cannot distinguish between these models and those which incorporate sharper conductivity contrasts.

THE MODEL MESHES

For a 2-D model the earth is parameterized by means of a grid of rectangular prisms, each having a uniform conductivity. The grid is terminated laterally by uniform layers and below by prisms elongated with depth. This grid is referred to as the regularization mesh. The individual blocks are made smaller than the data resolution length so that the locations of the block boundaries do not affect the final model. To perform the forward calculations required by the inversion scheme, a finite-element code described by Wannamaker et al. (1987) is used. This program uses a rectangular array of nodes to perform the finite-element calculations, which is called the finite-element mesh (see Figure 1). The finite-element mesh must contain at least the regularization mesh as a subset since there must be a node at every conductivity boundary. In normal practice, many nodes would be used across a conductivity boundary to ensure that the EM fields are computed correctly. However, the regularization mesh contains a much larger number of conductivity regions than is normally used for 2-D models based on assumed geologic structures, and so using several nodes for each conductivity element would be computationally expensive. The smooth inversion scheme will prevent large conductivity contrasts from appearing in the model, and several inversions have been conducted using no more nodes than defined by the regularization mesh. However, it is usually desirable to insert more nodes within the regularization mesh to maintain accuracy in the forward code, especially at the edges of the grid.

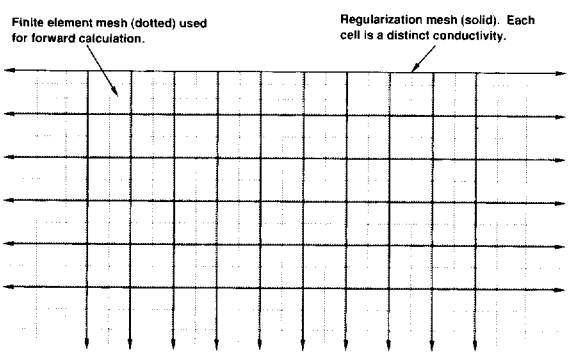

FIG. 1. The model is defined by the regularization mesh, which has a distinct conductivity value for each block. The forward computations require a finite element mesh which must at least include the boundaries defined by the regularization mesh, but for greater accuracy may include more nodes than this, especially near the surface and the edge and lower boundaries.

To maintain accuracy of the forward code, the spacing between nodes is taken to be approximately one-third of a skin depth, where the skin depth (δ) in a region of resistivity ρ ($\Omega \cdot$ m) is given in meters by

$$\delta \approx 500\sqrt{\rho/f}$$

and f is the frequency (Hz). The EM fields in a half-space decrease exponentially with depth, so the usual practice in MT surveys is to logarithmically space the frequencies at which the responses are calculated. Various depth scalings for 1-D smooth inversion were investigated by Smith and Booker (1988), and it is not surprising that a logarithmic depth scaling was found to fit the data most uniformly. Accordingly, a logarithmic depth scale is used for the node spacings and block sizes. The ideal depth scale is, of course, structure dependent and cannot be determined a priori for any real data.

Both the regularization and finite-element meshes remain fixed between iterations in this inversion method. To determine the sizes of the resistivity blocks, an estimate of the resistivity is made based on a 1-D inversion of the TE mode and inspection of the ρ_a curves. Since the inversion minimizes the model structure, the initial determination of the block sizes is not a critical step. The data generally require little conductivity contrast between the blocks, so it is not essential to have a fine grid in order to represent the models, and an excessively fine grid is, of course, no problem.

THE INVERSION METHOD

To suppress model structure not required by the data, the model roughness must be minimized. For a 2-D structure with x in the direction of the strike axis a measure of the model roughness may be given by

$$R_1 = \|\underline{\partial}_y \mathbf{m}\|^2 + \|\underline{\partial}_z \mathbf{m}\|^2, \quad (1)$$

where $\mathbf{m}$ is the vector of model parameters, $\underline{\partial}_y$ is a roughening matrix which differences the model parameters of laterally adjacent prisms, and $\underline{\partial}_z$ is a roughening matrix which differences the model parameters of vertically adjacent prisms. This is the expression for a first derivative roughness penalty. The penalty for the second derivative roughness is given by

$$R_2 = \|\underline{\partial}_y^2 \mathbf{m}\|^2 + \|\underline{\partial}_z^2 \mathbf{m}\|^2. \quad (2)$$

Since the model grid is terminated by uniform layers at the sides and uniform blocks below, first derivative smoothing best matches the boundary conditions imposed by the forward code. Therefore only the R_1 roughness penalty will be discussed. The vertical scale of the prisms is exponentially increased as a function of depth in order to coincide with the loss of resolving power; this is equivalent to increasing the penalty for roughness as a function of depth. The horizontal block boundaries and node spacings in the forward code extend to depth and are constrained by the requirement of having a fine mesh near the surface. Since the block widths

remain constant, the horizontal damping factors in the roughening matrix are adjusted to ensure that resolved features are not elongated in the vertical direction. That is, as the vertical scale of the prisms increases, a greater penalty must be given to differences in model parameters between laterally adjacent blocks.

Suppose the grid consists of a total of N elements, with p elements in the horizontal direction, each having width h, and ℓ elements in the vertical direction having widths v_i, $i = 1, 2, \ldots, \ell$. Numbering the elements from left to right starting at the top left element, the $N \times N$ vertical roughening matrix $\underline{\partial}_z$ is given by

$$\underline{\partial}_z = \begin{bmatrix} -1 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots \\ 0 & -1 & 0 & \cdots & 0 & 1 & 0 & \cdots \\ & & \ddots & & & & \ddots & \\ & & & -1 & & & \cdots & 1 \\ & & & \underline{0} & & & & \end{bmatrix}, \tag{3}$$

where $\underline{0}$ is a $p \times N$ matrix of zeroes. There are $p - 1$ zeroes between the entries in the rows of $\underline{\partial}_z$. Thus $\underline{\partial}_z$ acts to difference the model parameters between vertically adjacent blocks. The $N \times N$ horizontal roughening matrix $\underline{\partial}_y$ is given by

$$\underline{\partial}_y = \begin{bmatrix} \underline{\partial}_{y1} & & & \underline{0} \\ & \underline{\partial}_{y2} & & \\ & & \ddots & \\ \underline{0} & & & \underline{\partial}_{y\ell} \end{bmatrix} \tag{4}$$

where $\underline{\partial}_{yi}$ is the $p \times p$ horizontal roughening matrix for layer i, given by

$$\underline{\partial}_{yi} = \begin{bmatrix} -v_i/h & v_i/h & & \\ & -v_i/h & v_i/h & \underline{0} \\ & & \ddots & \\ & \underline{0} & -v_i/h & v_i/h \\ & & \cdots & 0 \end{bmatrix}. \tag{5}$$

Thus $\underline{\partial}_{yi}$ differences the model parameters between laterally adjacent blocks in layer i, basing the penalty for differences on the depth-to-width ratio of the blocks in that layer.

If the data are represented by $d_j, j = 1, 2, \ldots, M$, and one assumes that each of the data sets has known variances σ_j, a model's ability to fit the data can be quantified using the two-norm measure,

$$X^2 = \|\underline{\mathbf{W}}\mathbf{d} - \underline{\mathbf{W}}F[\mathbf{m}]\|^2, \tag{6}$$

where $F[\mathbf{m}]$ are nonlinear forward functionals acting upon the discretized model $\mathbf{m}$ to produce a model response and $\underline{\mathbf{W}}$ is the $M \times M$ diagonal weighting matrix

$$\underline{\mathbf{W}} = \text{diag}\{1/\sigma_1, 1/\sigma_2, \cdots 1/\sigma_M\}. \tag{7}$$

The smoothest model is sought subject to the criterion that it fit the data to a statistically reasonable tolerance. If it is assumed that the noise is uncorrelated and due to a zero-mean Gaussian process, then X^2 is chi-squared distributed with an expected value X^2_* equal to M, the number of independent data. That is, a model with an rms misfit of 1 is sought. The expected value is the best guess at the X^2 of the real Earth response. To be more conservative, one could choose to fit to 95 percent, or even 99 percent confidence limits, but since a larger number of data are used, there would be little difference in the resulting model. For example, for the 210–300 data items used in the examples, fitting to 95 percent confidence limits would require an rms error of 1.15 and fitting to 99 percent would require an rms error of 1.2. The models generated for 95 and 99 percent confidence limits would not be substantially different, and the use of the expected value does not require significantly more structure than the use of more conservative statistics. Furthermore, it is unusual to have field data with *errors* (not data) known accurately to within 15%, and so these arguments can only provide guidelines.

To solve the minimization problem, a Lagrange multiplier formulation is used and a stationary point is found for the unconstrained functional

$$U[\mathbf{m}] = \|\underline{\partial}_y \mathbf{m}\|^2 + \|\underline{\partial}_z \mathbf{m}\|^2 + \mu^{-1}\{\|\underline{\mathbf{W}}\mathbf{d} - \underline{\mathbf{W}}F(\mathbf{m})\|^2 - X^2_*\}, \tag{8}$$

where μ^{-1} is the Lagrange multiplier. This formulation resembles the regularization approach developed by Tihonov (1963a, b). It is common practice in regularization problems (e.g., Sasaki, 1989) to set the value of the Lagrange multiplier in advance and then solve for the model that best fits the data. However, this requires a priori knowledge of the model roughness and ignores the possibility of overfitting the data. Therefore, it is better to solve for a model that fits the data to within an acceptable tolerance. It is generally observed that as the tolerance is reduced, the model gets rougher.

The functional U is minimized at points where the gradient with respect to the model is zero. Since the data functionals are nonlinear, the functional U is linearized and solved iteratively. For a starting model $\mathbf{m}_1$, the first two terms of a Taylor expansion give the following approximation

$$F[\mathbf{m}_1 + \Delta] = F[\mathbf{m}_1] + \underline{\mathbf{J}}_1\Delta, \tag{9}$$

where $\underline{\mathbf{J}}_i$ is the Jacobean matrix, or an $M \times N$ matrix of partial derivatives of $F[\mathbf{m}_1]$ with respect to the model parameters, and

$$\Delta = \mathbf{m}_2 - \mathbf{m}_1 \tag{10}$$

is a small perturbation about a starting model. If these expressions are substituted back into equation (9), then the expression is obtained:

$$U = \|\underline{\partial}_y \mathbf{m}_2\|^2 + \|\underline{\partial}_z \mathbf{m}_2\|^2 + \mu^{-1}\{\|\underline{\mathbf{W}}\hat{\mathbf{d}}_1 - \underline{\mathbf{W}}\underline{\mathbf{J}}_1\mathbf{m}_2\|^2 - X^2_*\} \tag{11}$$

where

$$\hat{\mathbf{d}}_1 = \mathbf{d} - F[\mathbf{m}_1] + \mathbf{J}_1\mathbf{m}_1. \quad (12)$$

Note that U is now linear about $\mathbf{m}_2$. Differentiating with respect to $\mathbf{m}_2$ to find the model which minimizes U gives an iterative sequence for finding models;

$$\mathbf{m}_{i+1} = [\mu(\partial_y^T\partial_y + \partial_z^T\partial_z) + (\mathbf{W}\mathbf{J}_i)^T\mathbf{W}\mathbf{J}_i]^{-1}(\mathbf{W}\mathbf{J}_i)^T\mathbf{W}\hat{\mathbf{d}}_1. \quad (13)$$

A univariate search is conducted along μ at each iteration to find a model that minimizes the misfit to the data until the desired tolerance is obtained. The use of 1-D optimization at each iteration to choose the Lagrange multiplier is described in Constable et al. (1987), but it is instructive to extend their description and present the diagrams shown in Figure 2. An MT data set which may be fit conveniently with a 1-D model of only two layers is used. Fixing the thickness of the first layer and parameterizing in terms of the logs of the resistivities presents a problem that may be represented by contour diagrams of misfit versus the two parameters.

The inversion was started using a half-space of 1000 Ω · m and proceeded for four iterations before it converged. The upper frame of Figure 2 shows the whole arena for the inversion, while the lower frame is an enlargement of the region log $(\rho)_{1,2} = [-0.5, 1.5]$. The diagonal line on both frames is the set of maximally smooth models (i.e., $\rho_1 = \rho_2$) that includes the starting model (the triangle in the upper frame). The model chosen at the first iteration is shown in both frames as a circle. The solid line drawn to contain this point defines the set of all possible models attainable at the first iteration as the Lagrange multiplier sweeps from zero to infinity. Infinite Lagrange multipliers place all the weight on the smoothness condition, ignoring the fit to the data, and so the resulting perfectly smooth model lies on the diagonal, at about (0.6, 0.6) in this case. The other endpoint of the line corresponds to a Lagrange multiplier of zero, where all effort is made to fit the data regardless of model structure; this is the least-squares, or Gauss, step.

The starting model is well outside the region where linearization is a reasonable approximation (the roughly parabolic bowl in the lower frame), and the Gauss step is further from the solution than the starting point. The objective misfit was chosen to be 0.7, which cannot be achieved on the first iteration, so the algorithm chooses the model with the minimum misfit of 1.6. From this model, the second iteration can obtain the desired misfit (the model shown by the diamond). When the set of possible models intercepts the contour of desired misfit at two points, the intercept with the larger Lagrange multiplier corresponds to the smoother model. From now on the algorithm stays on the contour of desired misfit and moves to the point closest to the set of maximally smooth models. This is accomplished by iteration 3 (the asterisk), and verified by iteration 4, which does not change the model any further. Note that models 3 and 4 are close enough to the actual least-squares best fit that now the Gauss step does indeed yield its least-squares model.

Nearly all 1-D and many 2-D inversions are well behaved, converging with 8 to 12 forward calculations per iteration for the univariate optimization, with no modification to the steps required. However, occasionally the set of possible models does not include one with a lower misfit than the last model, or, once the desired misfit is acheived, one smoother than the last. This is due to a breakdown of the local linearity assumption or errors in the calculation of response functions and derivatives. The usual approach to safeguarding iterative least-squares solutions is to cut back the step size along the direction of the original step, by successively trying models with a step length of, say, 1/2, 1/4, 1/8 ... of the original. Rather than using this approach, in keeping with the

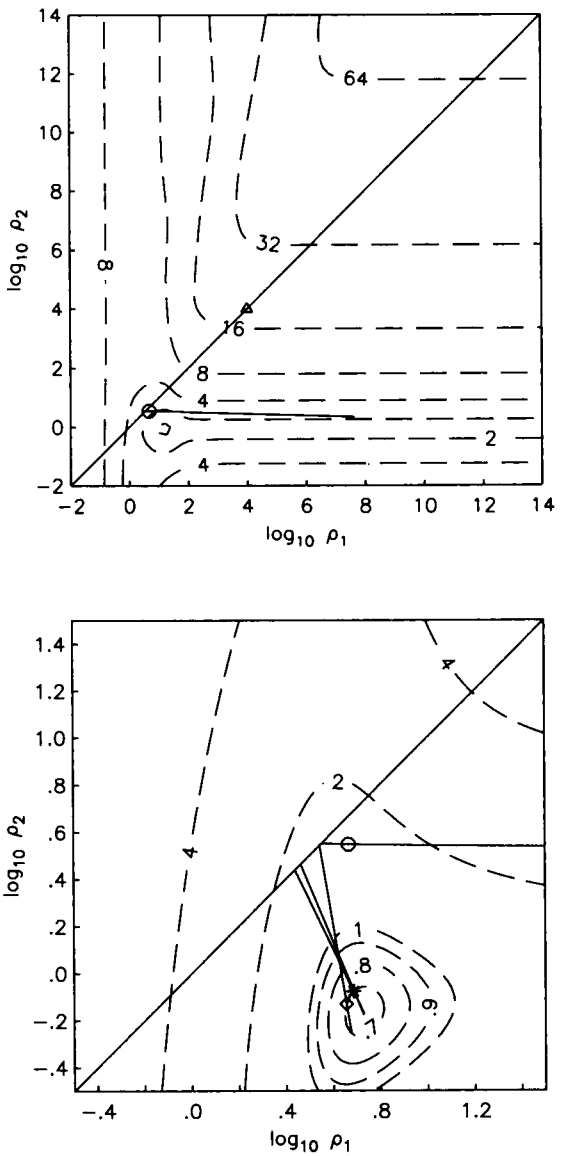

FIG. 2. Contours of the rms misfit in the $\log_{10}\rho_1$—$\log_{10}\rho_2$ plane. The lower frame is an enlargement of the area containing the minimum in the upper frame. The diagonal line on both frames is the set of maximally smooth models (i.e., $\rho_1 = \rho_2$). The starting model is shown by the triangle, the result of the first iteration by the circle, the second by the diamond, and the third and fourth both plot where the asterisk lies. The solid lines containing the models at each iteration show the sets of models attainable at each iteration as the Lagrange multiplier sweeps from zero to infinity. Infinite Lagrange multipliers correspond to Draconian smoothing and lie on the diagonal lines. Zero Lagrange multipliers correspond to the least-squares or Gauss steps, at the end points away from the diagonal.

algorithm for the original step, a new set of models is defined as a function of the Lagrange multiplier, each a distance of 1/2 (or 1/4, 1/8, etc.) between the present model and the set of models which has failed to provide an adequate step. That is, if the current model is $\mathbf{m}_i$ and the models which would normally be defined by the Lagrange multiplier are $F(\mu)$, then the new set of models is

$$G(\mu) = (1 - a)\mathbf{m}_i + aF(\mu),$$

where a is one for a normal step and is successively halved until an adequate step is found.

A flow chart indicating the steps used in the smooth inversion is shown in Figure 3. The method is completely general and may be applied to any nonlinear inverse problem. An accurate method of computing the forward problem and Jacobean matrix is essential to the implementation of the inversion scheme. The finite-element code for 2-D MT modeling, described in Wannamaker et al. (1987), is used here. The derivatives, found using a method described in Oristaglio and Worthington (1980), are also computed in this code.

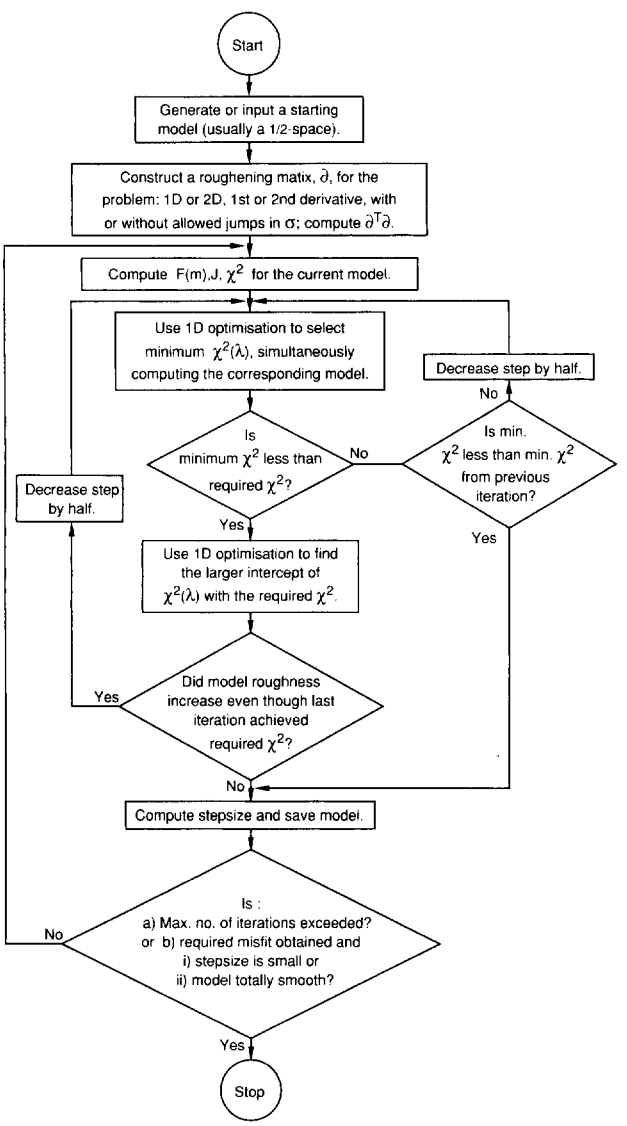

FIG. 3. A simplified flow chart of the inversion algorithm.

EXAMPLES FOR SYNTHETIC DATA

In the 2-D problem the electric and magnetic fields decouple into two modes: (1) the TE mode consisting of E_x, the component of the field parallel to the strike direction x, as well as B_y and B_z, the perpendicular components of the magnetic field; and (2) the TM mode consisting of B_x, E_y, and E_z. Based on forward modeling results, Wannamaker et al. (1984) argued that the response for a centrally located profile across an elongate 3-D body agrees with the TM response for a 2-D body with identical cross-section. They concluded that the 2-D interpretation of TE data anywhere in a 3-D setting or TM data near the edge of a 3-D anomaly would be erroneous because of the existence of current gathering. However, in the synthetic examples considered, the approximation to two dimensions is exact and the TE and TM modes are simultaneously inverted.

The smooth inversion method is first applied to a conductivity structure that consists of resistive and conductive prisms (2000 Ω · m and 5 Ω · m, respectively) embedded in a 100 Ω · m half-space. Data for both the TE and TM modes were generated at seven stations spaced at 10 km intervals, at eight periods ranging from 2.5 to 320 s. Figure 4a shows the model used to calculate the synthetic data. To simulate good quality, noisy data, two percent Gaussian noise was added to the data prior to inversion. The model is discretized into 392 resistivity blocks, with 14 rows of 28 blocks. A 30 Ω · m half-space was used as the starting model for the inversion.

Convergence was attained after 14 iterations, with one more iteration to verify convergence, yielding the model shown in Figure 4b. The parameters for the inversion in Figure 4b are given in Table 1 and the weighted residuals for every third station are shown in Figure 5. A resistivity contrast of a factor of 54 between the background and the conductive block is reached in the model, an overshoot of the original contrast of 20. However, a resistivity contrast of only 2.2 between the resistive block and the background was needed in order to fit the data. From the weighted residuals (Figure 5), it can be seen that the stations above the resistive structure are not underfit, so the lack of resolution for the resistive block is not an artifact of the inversion method.

Some authors (Oldenberg, 1988; Smith and Booker, 1990) have suggested that, due to charge buildup on the sides of structures with conductivity contrasts, TM data are more sensitive to resistive structure than are TE data, suggesting that a TM-only inversion would yield better lateral resolution than a mixed inversion. To test this, separate TE and TM inversions for the model shown in Figure 4a indicate that, while the resistive structure is imaged better for TM only than for TE only, the maximum resistivity contrast for the resistive block in the TM-only inversion is 1.7, and so the lateral resolution is in fact worse for TM-only than for the joint TE-TM inversion. Since the TM responses are spatially narrower than the TE responses, the number of stations was increased to one station every 5 km and data with 2 percent Gaussian noise were generated at the same frequencies as

before. Subsequent TM-only inversion showed only a slight improvement over the previous TM inversion for stations every 10 km, with the resistivity contrast for the resistive block increased to 1.8. It may be concluded that resistive structures whose size is comparable to the depth of burial can be only poorly resolved with MT data.

Smith and Booker (1988) discuss the use of Spearman's statistic to rule out systematic bias in fit as a function of frequency. Trends in the data residuals might indicate that (1) the assumption about the dimensionality of the region is incorrect, (2) the data errors are incorrectly estimated, or (3) there is a systematic bias in the inversion method. For synthetic data, the first two possibilities obviously can be ruled out. Smith and Booker call a fit in which the residuals are evenly distributed over all frequencies "white" and discuss how overfitting any particular frequency band implies that noise is being fit while underfitting data in a frequency band leads to the loss of significant structure. The plots in Figure 5 show that the weighted residuals in the model responses are essentially uncorrelated with frequency and TE and TM modes are fit equally well. Thus it appears that using exponential depth scaling and horizontal weight-

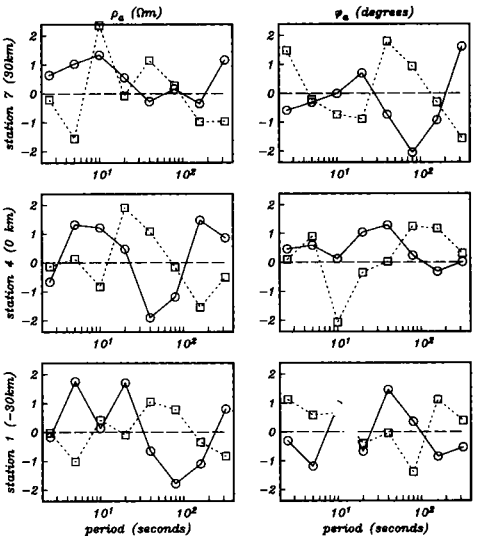

FIG. 5. Normalized residuals for the model shown in Figure 4 for the middle and end stations. The circles and squares indicate the TE and TM residuals, respectively, as a function of period.

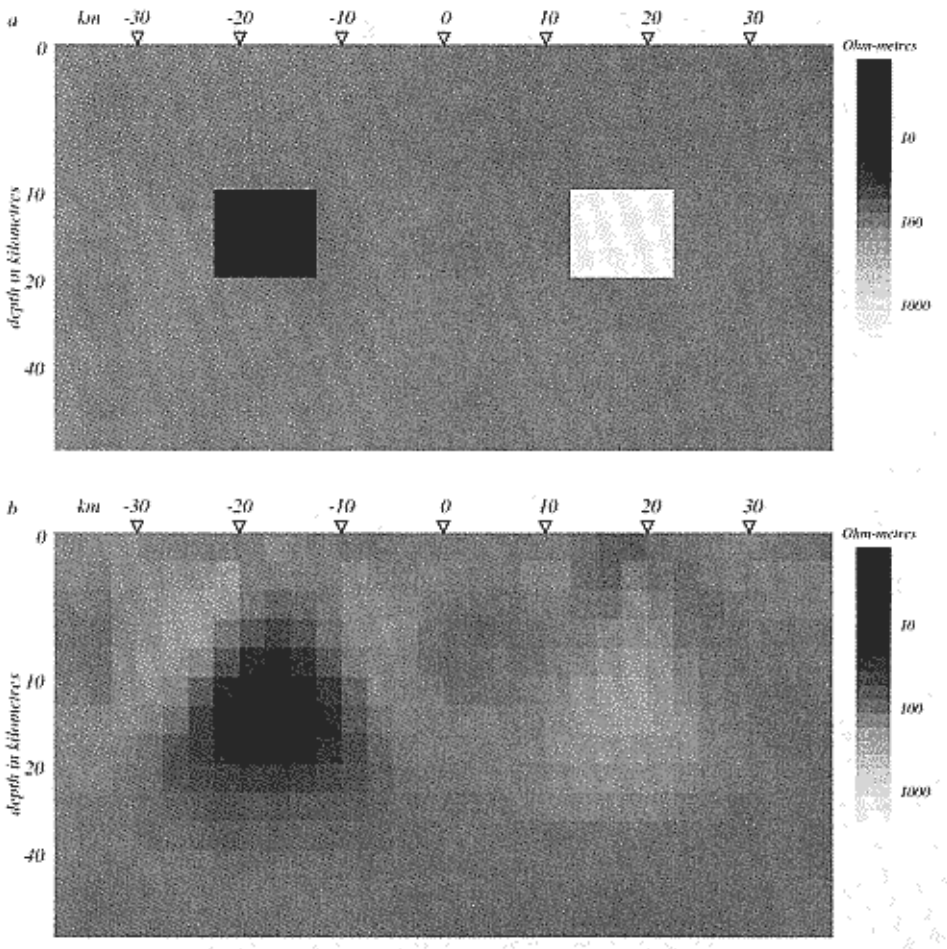

FIG. 4. (a) The model used to generate the data for the inversion. It consists of a high-conductivity block (5 Ω · m) and a low-conductivity block (2000 Ω · m in a 100 Ω · m) half-space. Two percent Gaussian noise was added to the data prior to inversion. (b) The results of the inversion after 14 iterations.

ing factors effectively eliminates preferentially fitting subsets of the data.

The CPU time used for for the model in Figure 4a was approximately 11 hours per iteration on a Mac II computer and required 6.4Mb of memory. On a CRAY X-MP the time per iteration is about 2 minutes. Approximately 80 percent of the CPU time is spent on partial derivatives, the forward calculations use about 15 percent of the time, with the remaining 5 percent for the inversion code. As shown in Table 1, about seven forward calculations are required for each iteration. Since each forward calculation is 30 times faster than doing the partial derivatives for the whole model, completing a number of forward calculations to find the minimum in U at each iteration is an efficient means of convergence. As the model size grows, the matrix inversion will begin to dominate the CPU time. This would be especially true if the size of the inversion were increased by adding stations (rather than frequencies), which would add very little to the forward model calculations but burden the inversion. Cholesky decomposition is used, which is not the fastest method of inverting a positive definite symmetric matrix but is very stable.

Table 1. Parameters for the inversion shown in Figure 4. Asterisks indicate that the step size was reduced to ensure convergence. The number of forward calculations performed at each iteration is given by N.

Iteration	rms	R	Step size	$\log_{10}(\mu)$	N
0	23.1	0.0			
1	17.1	75.7	503	1.16	7
2	12.0	1×10^{-8}	644	8.81	10
3	6.56	.52	99.1	3.21	10
4	4.02	2.85	22.7	2.21	6
5	2.68	4.75	3.39	2.21	6
6	2.23	5.48	2.45	2.21	6
7	1.33	7.41	3.02	1.59	7
8	1.28	6.67	0.59	1.75	7
9	1.058	7.58	0.52	1.35	6
10	1.055	7.48	0.20	1.35	7
11	1.01	7.96	0.24	1.15	6
12*	1.00	7.43	0.06	1.50	14
13	1.00	7.10	0.11	1.53	7
14	1.00	7.05	.016	1.57	7
15	1.00	7.03	9.8×10^{-4}	1.58	7

It is possible that a number of minima in the functional U exist, leading to cases in which the final model depends upon the starting model. This cannot be ruled out on the basis of trial and error using different starting models because there always may be some other, untried, starting model which could lead to a different solution. If convergence to the same model is obtained starting from a very rough model and starting from a homogeneous half-space, one would have more confidence that the globally smoothest model had been found. According to the objectives of this inversion scheme, any unnecessary structure should be smoothed out of any model fitting the data. To test the approach, the true model was used as a starting point and after six iterations the same model as before was attained.

A priori information about the location of discontinuities is a powerful tool in resolving structure. If it is known where sharp discontinuities in resistivity exist (e.g., the base of a sedimentary basin as determined by seismic information), then the penalty for roughness at those boundaries may be removed by zeroing the entries in the roughening matrices which correspond to the sharp boundaries. It should be stressed that structure may still appear at locations other than the known boundaries. As an example of this technique, the penalty for structure at the outside boundaries of each zone was removed in the model containing high- and low-resistivity zones. The model after three iterations in an inversion of the same data as before is shown in Figure 6. The initial model is very nearly recovered and the resistive zone is now well-defined. Because insertion of a sharp boundary changes the resolving power so dramatically, caution must be exercised when using this technique. Incorrect placement of boundaries may produce misleading results.

The next example shows the result of inverting data generated by a conductive sill (5 $\Omega \cdot$ m) terminating short of an equally conducting body and embedded in a more resistive (100 $\Omega \cdot$ m) host as shown in Figure 7a. A geologic

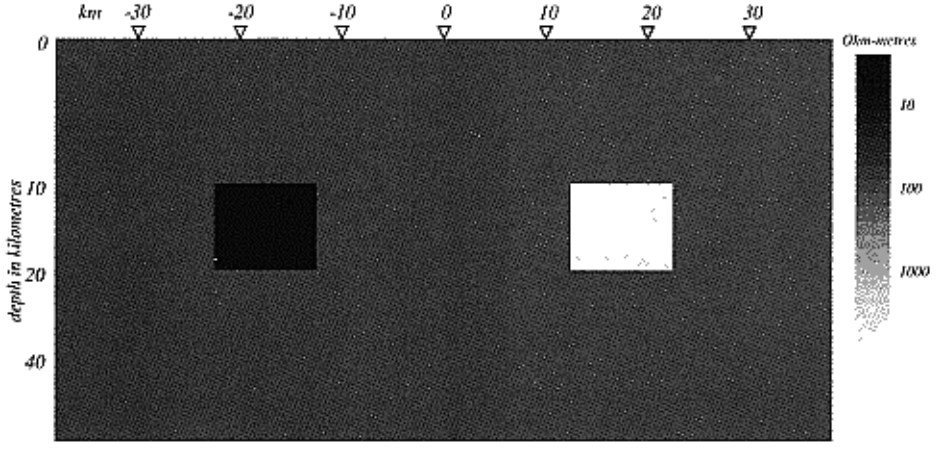

FIG. 6. The maximally smooth model produced when the penalty for structure at known discontinuities is removed. The model discretization and data used for the inversion were the same as for the inversion in Figure 4.

analog of this example might be a conductive subducting slab terminating short of a conductive basement. A structure of this type may cause difficulty in approximate inversion techniques since, for some stations, the data are more strongly dependent upon the conductivity of the blocks in the resistive gap than on the blocks directly beneath the station. Seven stations were placed at a spacing of 10 km, and both TE and TM mode responses were calculated at each site for nine periods ranging from 4.5 s to 2128 s. Two percent Gaussian noise was added to the data set prior to inversion and a 50 $\Omega \cdot$ m half-space was used as the starting model for the inversion.

Convergence to the model shown in Figure 7b was attained at the thirteenth iteration. The high-resistivity blocks near -10 km and $+10$ km could be caused by fitting noise or instabilities in the forward model. These effects are not seen in the previous example (Figure 4), and it is concluded that these may be due to Gibb's phenomenon. That is, in trying to fit a smooth model to a feature exhibiting sharp discontinuities an overshoot in the fit is obtained. Such behavior is very common in 1-D model studies. Inspection of the weighted residuals (not shown here) again indicates that there is no systematic bias in the fit. The model is still quite rough at a depth of 10 km, which indicates that the result is somewhat dependent upon the model parameterization but a finer grid of blocks would result in a smoother model. The resistive region separating the conductive areas is well resolved, indicating that the inversion algorithm is robust to a high degree of nonuniformity in the Jacobean matrix. This simple example suggests that a resistive gap between two conducting areas may be resolvable using 2-D MT data.

A priori information about the conductivity of a given area may also improve the ability to resolve structure in the remainder of the model. For example, if the shape and conductivity of a given structure are known, then one may include this information in the model and invert for the remainder of the model. Such a situation may arise if data are gathered near a coastline, where the conductivity and bathymetry are well known, or if the data are collected over a sedimentary basin in which the layer thicknesses and resistivities are known from seismic studies and well-logging information. As an example of this technique, data were

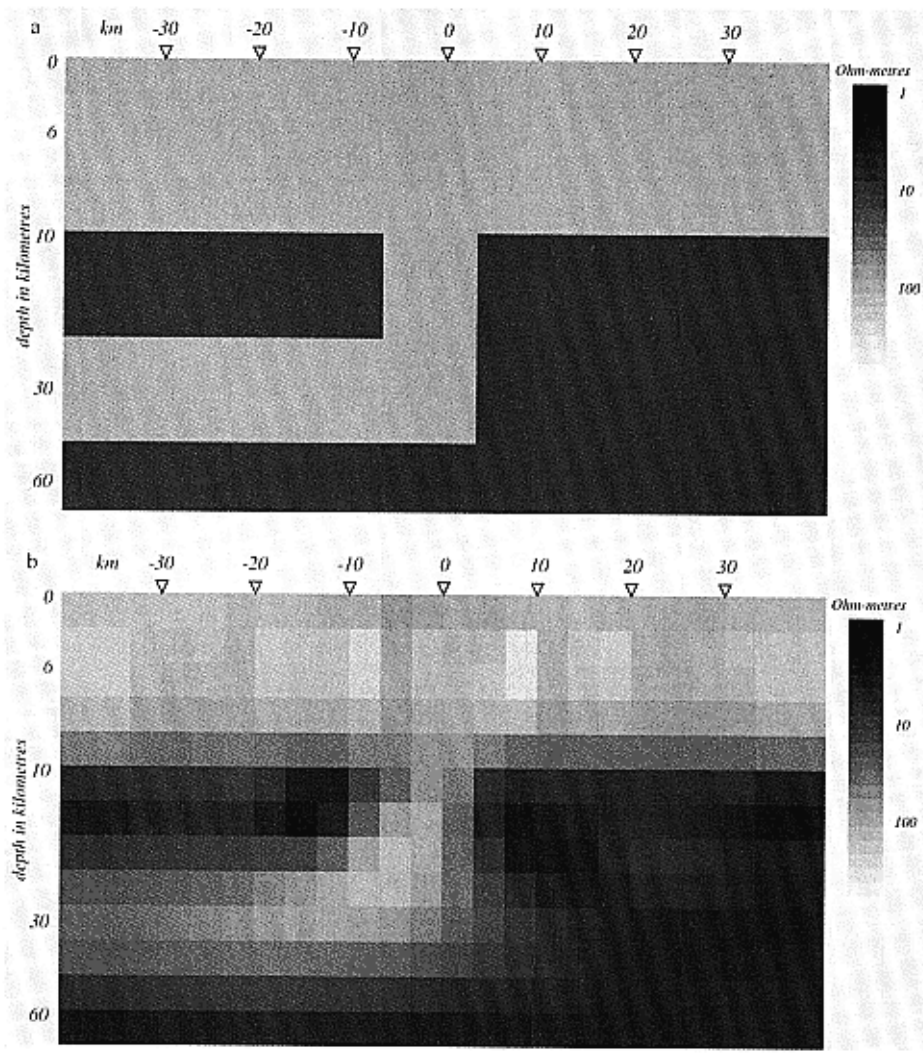

FIG. 7. (a) The model used to generate the data for the inversion consists of a conductive ledge terminating short of a conducting body. (b) The model obtained after inversion.

generated from the model shown in Figure 8a. The model represents a sedimentary basin consisting of three layers having resistivities of 80, 200, and 50 Ω · m from top to bottom. Below the sedimentary basin is an anticlinal structure having a resistivity of 50 Ω · m surrounding a structure having a resistivity of 500 Ω · m. These structures are embedded in a half-space having a resistivity of 1000 Ω · m. The data, with five percent Gaussian noise added, were generated at eight periods ranging from 0.01 to 33 s at 11 stations placed 2 km apart.

The data were inverted twice, once with the resistivity of the sedimentary basin held fixed and once with no a priori information. A 100 Ω · m half-space was used for the starting model in the second inversion while the known resistivities of the sedimentary basin were embedded in a 100 Ω · m half-space for the starting model in the first inversion. Convergence to the model shown in Figure 8b was attained after 10 iterations for the case where the resistivity of the sedimentary basin is known; convergence to the model shown in Figure 8c was attained after 14 iterations. The sedimentary basin shows up very poorly in the case where the known structure is not included in the model. There is a marginal improvement in the resolution of the anticlinal structure in the case where the known conductivity of the sedimentary basin is included.

INVERSION OF FIELD DATA

The 2-D inversion of field data introduces problems associated with computation and breakdown of assumptions

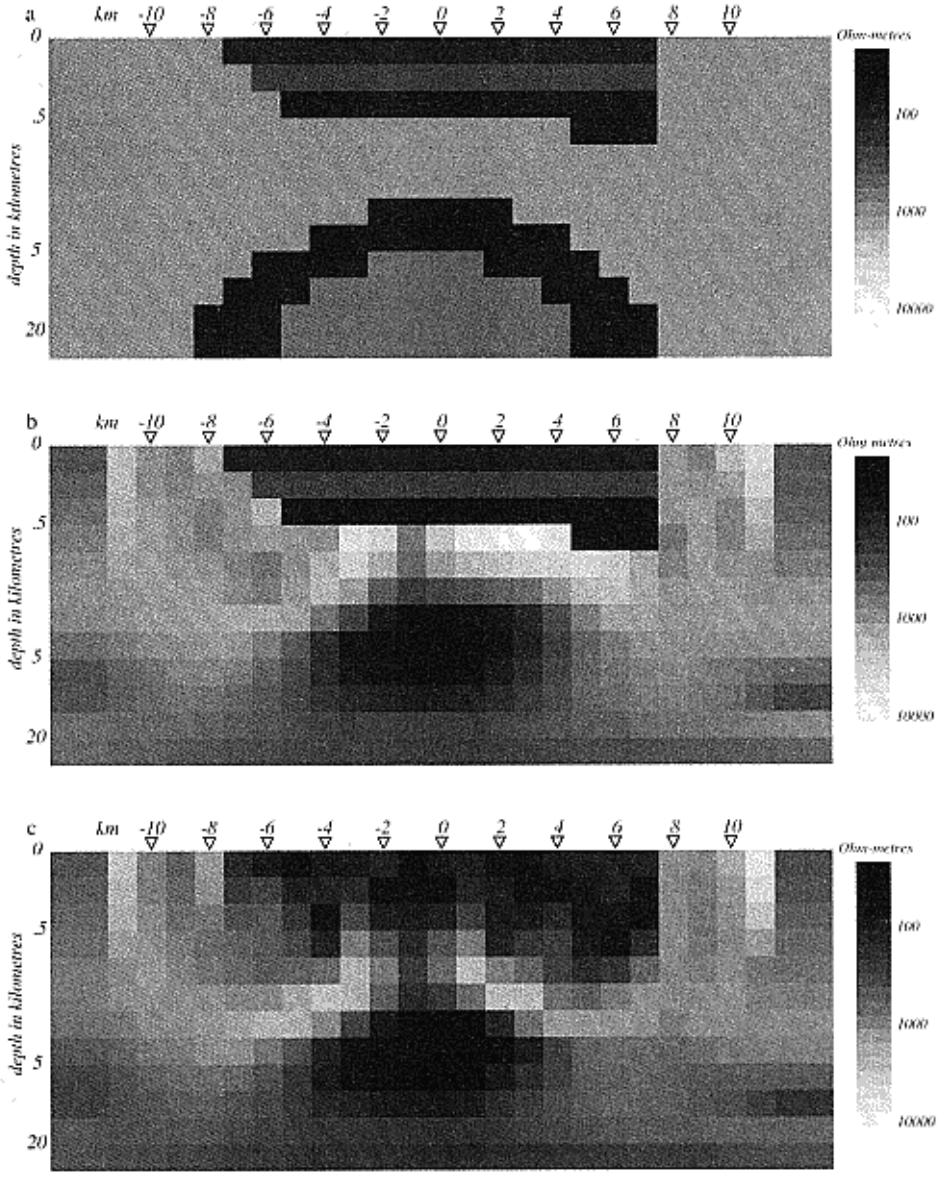

FIG. 8. (a) The model used to generate the data for the inversion represents a sedimentary basin over an anticlinal structure. (b) The model obtained when the conductivity of the sedimentary basin is held fixed. (c) The model obtained for the case of no a priori information.

about dimensionality and noise. Computational difficulties include: (1) the amount of memory required for handling broadband data amount due to the amount of data, and (2) the large variation of skin depths involved. The skin depth of the highest frequencies require the resistivity blocks to be very small near the surface while the lower frequencies require the model to extend to large depths. Since the block widths propagate to depth, a very large mesh is required for the inversion. Another problem is that inversion of 2-D data is much slower than the 1-D case. Since the required time for the forward calculation increases linearly with the number of frequencies used, if the original data are oversampled in frequency, the computation time can be reduced by using a subset of the data for the initial iterations and adding more data as convergence is attained.

The nature of the noise and noise estimates also presents difficulties. The two-norm measure assumes Gaussian errors in the data and is not robust to the presence of outliers which result from non-Gaussian noise. Also, if the actual structure is 3-D, then the data may be very accurate and still not correspond to any 2-D solution. The question of existence of a solution for any 1-D MT data set has been solved (Parker, 1980; Parker and Whaler, 1981). However, there is not as yet a way of knowing whether any 2-D structure will fit a given data set, or, if so, what the minimum possible misfit would be.

A subset of the COPROD2 data, provided by Alan Jones of the Geological Survey of Canada, was used as an example for the inversion of field data. These data were collected over the central plains anomaly in the province of Saskatchewan, Canada, and have been corrected for static shift (Jones, 1988). The survey line was east-west, with receiver spacings of approximately 10 km. The TE and TM responses appear quite uniform over the high frequencies up to 4 s (Jones and Savage, 1986), suggesting that the surface structure is 1-D to a first approximation. At lower frequencies the TE and TM responses diverge, indicating higher dimensional structure at greater depth. For this reason, four of the lower frequencies were chosen at 20 stations at the eastern end of the survey area as a subset for inversion. Since regional structure was the target for interpretation, relatively large blocks having widths of 10 km were used and stations exhibiting large conductivity contrasts with its neighbours were eliminated. The inversion presented here is meant only to show the application of the 2-D MT inversion to real data and does not represent a full interpretation of the COPROD2 data set (to follow in a subsequent paper).

If it can be assumed that the Earth structure is largely 2-D in this area, 3-D structure will appear in the data as noise. Some of the data used in the inversion had very small errors which the authors felt were unrealistic for the 2-D approximation, so the minimum error was increased to 10 percent. Starting with a half-space of 100 $\Omega \cdot$ m, convergence to an rms misfit of 1.0 was acheived after 28 iterations. The resulting model is shown in Figure 9. Two conductivity anomalies which have been identified in previous work (Jones and Craven, 1989) are indicated here: (1) the North American Central Plains (NACP) anomaly below the center of the array at a depth of 6–25 km, and (2) the shallower Thompson Belt anomaly to the east of the survey area. The NACP anomaly appears to be very discontinuous, possibly due to static shifts remaining in the data, which often have the effect of introducing vertical structure (Jones, 1988). This suggests that one might be able to remove static shift by introducing a shift parameter at each station and simultaneously solving for the smoothest model.

The data and model responses are plotted as a function of station location in Figure 10. A non-Gaussian error distribution would be manifested as non-Gaussian residuals; however, a Kolmogorov-Smirnoff test on the residuals cannot exclude the Gaussian hypothesis with any significant confidence. This indicates that the misfit is not likely due to 3-D structure, since in that event large outliers in the residuals would be expected. Since the inversion algorithm, which performs well on 1-D and 2-D synthetic data, could not

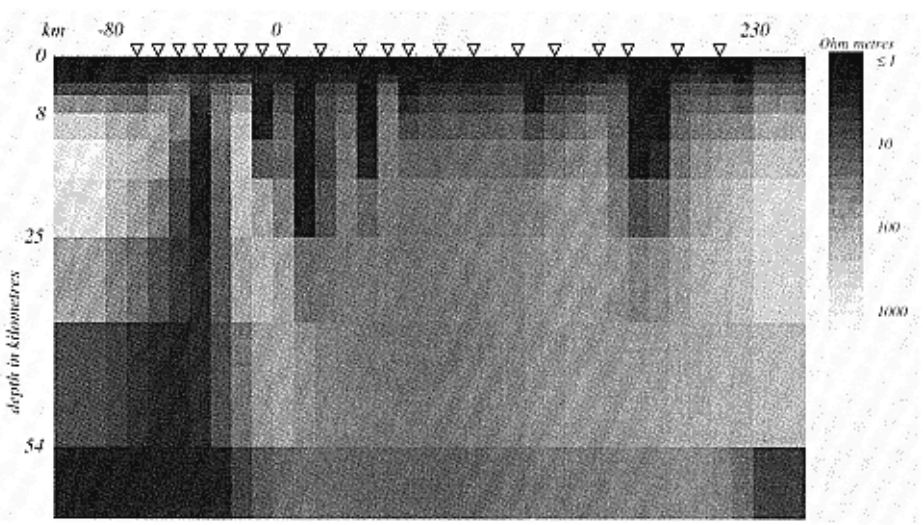

FIG. 9. Inversion of a subset of the COPROD2 data with minimum 10 percent error. Two known anomalies are defined, the NACP anomaly at the center of the survey line and the Thompson Belt anomaly indicated at the left of the model. A linear depth scale is used here to more clearly indicate the geologic structure.

achieve a misfit of rms = 1 with the original data set, enlarging the errors in the data may be justified.

DISCUSSION AND SUMMARY

We emphasize that there is no guarantee that a 2-D model exists which adequately fits a given data set. The issues of existence and adequate fit depend on the choice of model dimension and parameterization and on the reliability of data errors. The model complexity must be adequate to represent the real complexity of the Earth, and the data errors must be well estimated. The number of dimensions of the model must be at least as large as the regionally significant dimensionality of the Earth. In the context of this paper, it is assumed that the Earth may be adequately represented by a 2-D model. There undoubtedly exists 3-D structure, but a 2-D model is adequate as long as the expression of this structure is within the estimated data error. It is obvious that the concept of an adequate fit depends greatly on the data errors. If the data errors are systematically underestimated, then it will be difficult or impossible to find a model which fits the data adequately. If the random scatter in the data exceeds the effect of 3-D structure, then 2-D modeling will allow one to fit the data, so long as the data errors have not been underestimated.

An MT inversion algorithm was developed to solve for the smoothest 2-D model fitting a given data set. The resistivities of a large number of blocks are determined iteratively, with convergence generally being attained in 10 to 20 iterations. While it cannot be proven that a global minimum in model roughness has been reached, tests on a simple model indicate convergence to the same solution is reached from both a starting model which is rough and from a half-space starting model. Unlike most regularized inversion algorithms, the value of the trade-off parameter between data misfit and model roughness is not determined in advance. This gives greater flexibility in fitting the data to a statistically reasonable tolerance while simultaneously keeping the roughness to a minimum.

There is some similarity between this method and that of Smith and Booker (1990). Note, however, that these authors make several approximations in order to realize large improvements in computation time. A comparison of the resistive block model (Figure 4) suggests that their approximations result in models which are slightly rougher than necessary, although the models are generally in good agreement. However, even if the approximate methods were to yield identical results, the more exact method needs to be available before confidence in the faster algorithm is possible. The reader should bear in mind that the models compared to date are very simple. In fact, Smith and Booker (1990) do not present inversions of real data and have restricted themselves to inverting for simple prismatic structures.

Although computationally expensive, the algorithm presented here is versatile, allowing extension to methods other than MT, such as resistivity, and incorporation of varying levels of a priori information. The most objective initial approach to using this method is, of course, to include no a priori information, allowing one to obtain an idea of the resolving power of the data. If the location of sharp discontinuities in resistivity is known, the penalty for roughness at the boundaries may be removed. Finally, if both the resistivity and boundary locations are known in advance, these may be specified and will not be updated in further iterations. Because the addition of incorrect a priori information produces misleading results, care must be used when incorporating "known" structure.

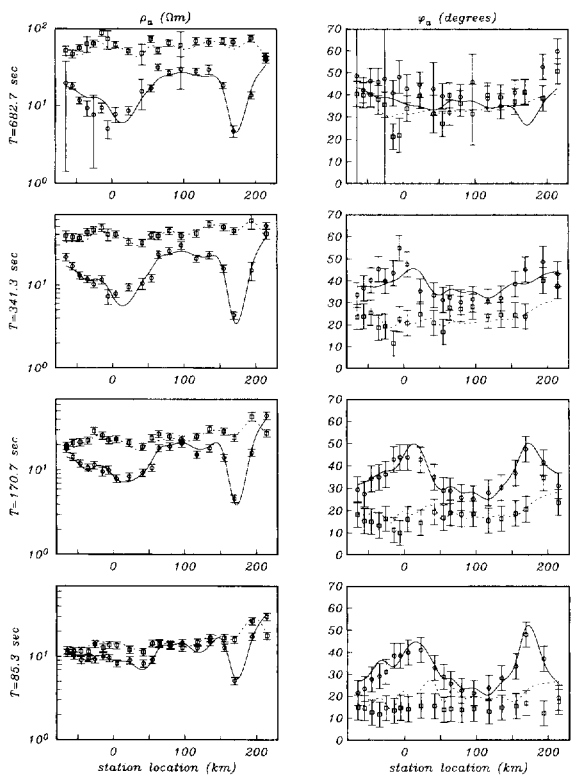

FIG. 10. Data and model responses plotted as a function of station location. The circles and squares indicate the TE and TM data, respectively, while the solid and dashed lines respectively indicate the TE and TM responses. Negative TM phases are plotted.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to Phil Wannamaker for providing his 2-D MT code, without which this work could not have proceeded, and for providing advice on its use. They also thank Alan Jones for making the COPROD2 data available and for providing help with and computer time for the COPROD2 inversions. Finally, the authors owe a great debt to Bob Parker for originating the smooth inversion and his continued interest in this work. This work was supported by DOE contract DE-FG03-87ER13779.

REFERENCES

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

EMSLAB Group, 1988, The EMSLAB electromagnetic sounding experiment: EOS Trans., Am. Geophys. Union, **69**, 98–99.

Jones, A. G., 1988, Static shift of magnetotelluric data and its removal in a sedimentary basin. Geophysics, **53**, 967–978.

Jones, A. G., and Craven, J. A., 1989, The North American Central Plains anomaly and its correlation with gravity, magnetic, seismic, and heat flow data in Saskatchewan, Canada: Phys. Earth Planet. Inter., **53**, 967–978.

Jones, A. G., and Savage, P. J., 1986, North American Central Plains conductivity anomaly goes east: Geophys. Res. Lett., **13**, 685–688.

Jupp, D. L. B., and Vozoff, K., 1977, Two-dimensional magnetotelluric inversion: Geophys. J. Roy. Astr. Soc., **50**, 333–352.

Oldenburg, D. W., 1988, Inversion of electromagnetic data: an overview of new techniques: ninth workshop on electromagnetic induction in the Earth and Moon.

Oristaglio, M. L., and Worthington, M. H., 1980, Inversion of surface and borehole electromagnetic data for two-dimensional electrical conductivity models: Geophys. Prosp., **28**, 633–657.

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

Parker, R. L., and Whaler, K. A., 1981, Numerical methods for establishing solutions to the inverse problem of electromagnetic induction: J. Geophys. Res., **86**, 9574–9584.

Rodi, W. L., Swanger, H. J., and Minster, J. B., 1984, ESP/MT: an interactive system for two-dimensional magnetotelluric inversion: Geophysics, **49**, 611.

Sasaki, Y., 1989, Two-dimensional joint inversion of magnetotelluric and dipole-dipole resistivity data: Geophysics, **54**, 254–262.

Smith, J. T., and Booker, J. R., 1988, Magnetotelluric inversion for minimum structure: Geophysics, **53**, 1565–1576.

Smith, J. T., and Booker, J. R., 1990, The rapid relaxation inverse for two- and three-dimensional magnetotelluric data: J. Geophys. Res., submitted.

Tihonov, 1963a, Solution of incorrectly formulated problems and the regularization method: Soviet Math. Dokl., **4**, 1035–1038.

——— 1963b, Regularization of incorrectly posed problems: Soviet Math. Dokl., **4**, 1624–1627.

Wannamaker, P. E., Hohmann, G. W., and Ward, S. H., 1984, Magnetotelluric responses of three-dimensional bodies in layered earths: Geophysics, **49**, 1517–1533.

Wannamaker, P. E., Stodt, J. A., and Rijo, L., 1987, A stable finite-element solution for two-dimensional magnetotelluric modeling: Geophys. J. Roy. Astr. Soc., **88**, 277–296.