Inversion of induced polarization data

Douglas W. Oldenburg* and Yaoguo Li*

ABSTRACT

We develop three methods to invert induced polarization (IP) data. The foundation for our algorithms is an assumption that the ultimate effect of chargeability is to alter the effective conductivity when current is applied. This assumption, which was first put forth by Siegel and has been routinely adopted in the literature, permits the IP responses to be numerically modeled by carrying out two forward modelings using a DC resistivity algorithm. The intimate connection between DC and IP data means that inversion of IP data is a two-step process. First, the DC potentials are inverted to recover a background conductivity. The distribution of chargeability can then be found by using any one of the three following techniques: (1) linearizing the IP data equation and solving a linear inverse problem, (2) manipulating the conductivities obtained after performing two DC resistivity inversions, and (3)

solving a nonlinear inverse problem. Our procedure for performing the inversion is to divide the earth into rectangular prisms and to assume that the conductivity σ and chargeability η are constant in each cell. To emulate complicated earth structure we allow many cells, usually far more than there are data. The inverse problem, which has many solutions, is then solved as a problem in optimization theory. A model objective function is designed, and a "model" (either the distribution of σ or η) is sought that minimizes the objective function subject to adequately fitting the data. Generalized subspace methodologies are used to solve both inverse problems, and positivity constraints are included. The IP inversion procedures we design are generic and can be applied to 1-D, 2-D, or 3-D earth models and with any configuration of current and potential electrodes. We illustrate our methods by inverting synthetic DC/IP data taken over a 2-D earth structure and by inverting dipole-dipole data taken in Quebec.

INTRODUCTION

Induced Polarization (IP) data are routinely collected in mineral exploration surveys and they are also finding their niche in environmental surveys (Barker, 1990). Excellent reviews on the IP method and case histories can be found in Sumner (1976), Bertin and Loeb (1976), Fink et al. (1990), and Ward (1990). The difficulty with interpreting IP data is the lack of flexible, efficient, and robust inversion algorithms. This is reflected in the remark in Hohmann (1990) regarding the availability of forward algorithms and the lack of general inversion techniques for the IP method. Consequently, much of today's interpretation is still carried out by working with pseudosections. Yet only in very simplistic circumstances will the images on the pseudosections emulate geologic structure, and consequently, inferences made about the substructure directly from the data are often

incorrect. As an illustration, we present a fairly simple example of DC/IP data taken over a 2-D earth structure. The true conductivity and chargeability models and the pseudosection plots obtained by carrying out a pole-pole DC/IP survey are shown in Figure 1. Except for the region near the surface, there is little compelling evidence in the IP pseudosection to indicate a chargeable body at depth.

Methods for inverting IP data do exist but literature on this subject is sparse. In defining IP data, most authors adopt a presentation given in Siegel (1959) that the ultimate effect of a chargeable body is to alter its effective conductivity. As such, the IP and DC resistivity problems are intimately linked, and the inversion of IP data is a two-step process. In the first stage, the DC potentials are inverted to recover the background conductivity σ_b. The second step accepts σ_b as the true conductivity of the medium and attempts to find a chargeability that satisfies the data. This is usually done by

linearizing the equations about σ_b to produce a system of equations that can be solved for the chargeability distribution.

Early algorithms for inverting DC and IP data generally parameterized the earth model into a relatively small number of blocks and kept the same parameterization for inverting the DC and IP data (Pelton et al., 1978; Sasaki, 1982; Rijo, 1984). Overdetermined systems of equations were solved and algorithm convergence was judged on the basis of data misfit alone. There are practical difficulties that arise with this approach. The electrical conductivity structure of the earth is complicated and rarely does a representation by a few blocks adequately represent the true distribution of this physical property. Also, anomalous regions of high or low conductivity do not necessarily correspond to regions of high chargeability. If only a few blocks are used, keeping the same parameterization may preclude the possibility of finding a meaningful solution. In addition, the restriction of using only a few model cells does not allow insight about the nonuniqueness that is inherent in the inverse problem. The difficulties with respect to parameterization can be overcome by discretizing the earth into a large number of cells. The inverse problem is then solved as an optimization problem where an objective function of the model is minimized subject to adequately fitting the data. An archetypal example applied to a geophysical inverse problem was presented by Constable et al. (1987). LaBrecque (1991) has applied this methodology to carry out a 2-D inversion of IP data in a

cross-borehole tomography experiment, and Beard and Hohmann (1992) have provided an approximate inversion of IP data that is valid when resistivity contrasts are small.

In this paper, we use Siegel's (1959) formulation and develop three methods by which to invert IP data. In the first method, we assume that the chargeability is small and linearize the equations. The technique is similar to that presented in LaBrecque (1991), but we use a more general model objective function, we incorporate a subspace methodology to bypass the large computations normally required to invert the full matrix system, and we work with chargeabilities directly in the inverse problem rather than with their logarithms. The second method makes use of formal mappings connecting DC/IP voltages with a conductive and chargeable earth. Two DC resistivity inversions of different data are carried out, and the chargeability is obtained by manipulating the recovered conductivities. The third method makes no assumptions about the size of η and solves a nonlinear inverse problem to recover the chargeabilities. The resultant algorithm is essentially the same as that used to invert the DC resistivity data. All of our techniques are applicable to any dimension of earth structure and to any configuration of electrodes. Our paper begins by defining the forward mapping for the IP data. Next, we introduce the 2-D synthetic example and invert pole-pole DC potentials to recover the background conductivity. This is presented in some detail because the subspace method used to invert the DC data also plays the dominant role in inverting IP data.
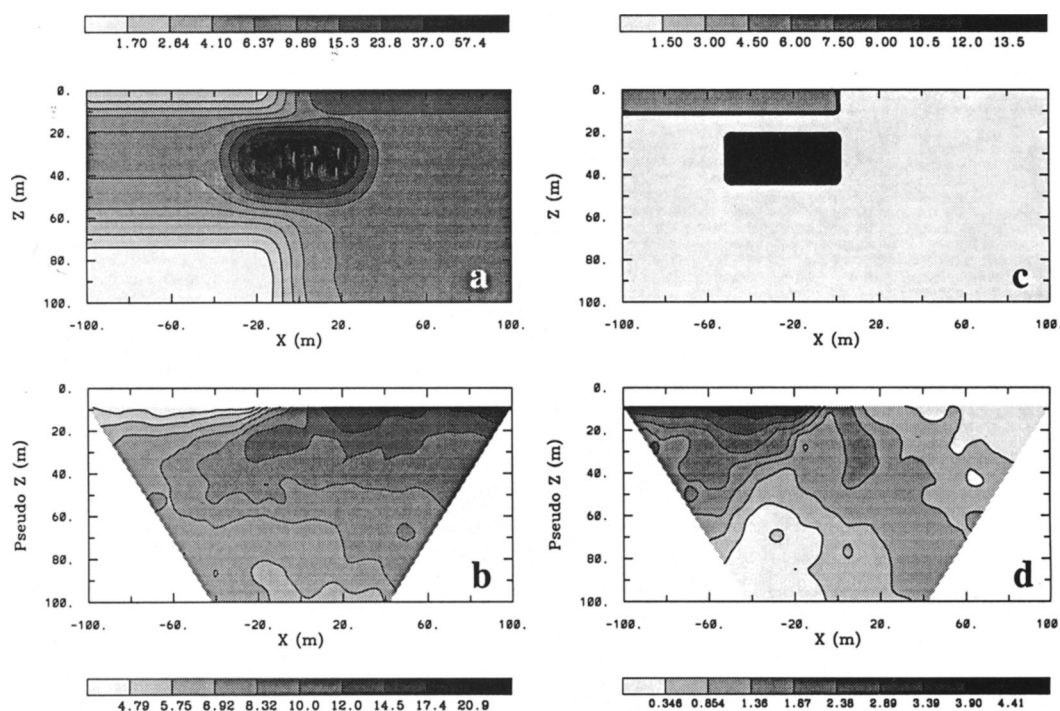

FIG. 1. A synthetic 2-D conductivity model is shown in (a). Surface electrodes are spaced at 10 m over the interval $x = (-100, 100)$ m. Current is input at each electrode site in turn, and potentials are observed at the remaining 20 sites. The pole-pole potentials ϕ_σ are converted to apparent conductivity $\sigma_a = I/(2\pi r\phi_\sigma)$, where r denotes the distance between current and potential electrodes and are plotted in (b). The apparent conductivity value is plotted midway between the current and potential electrodes at a (pseudo) depth of $z = 0.86r$. The grey scales indicate conductivity in mS/m. The synthetic chargeability model is given in (c) and the apparent chargeability, plotted in the pseudosection format is shown in (d). The grey scale indicates chargeability in percent.

Three methods for inverting IP data are explained and illustrated with the synthetic example. We address the issues of positivity, of how well the background conductivity needs to be known, and practical considerations pertaining to the data to be inverted. We also examine the merits of using η or ln η as the variable in the inversion and show how minimizing different objective functions can be useful in hypothesis testing and in exploring nonuniqueness. A field data set of dipole-dipole data is inverted, and the paper concludes with a discussion about the relative merits of the inversion algorithms.

FORWARD MODEL

Complete understanding of the microscopic phenomenon that result in the macroscopic IP response has not been achieved. Here we adopt a macroscopic representation of the physical property governing the IP response that was put forth in Siegel (1959). Basically, he introduces a macroscopic physical parameter called chargeability to represent all of the microscopic phenomena. As such, our earth model is described by the two quantities: conductivity $\sigma(x, y, z)$ and chargeability $\eta(x, y, z)$. Both are positive, but while conductivity varies over many orders of magnitude, chargeability is confined to the region [0, 1). We note that Siegel's model refers only to the volumetrically distributed polarization and does not apply to highly conductive targets in which surface polarization dominates the IP effect. Fortunately, this is not an important limitation for the majority of practical situations.

A typical IP experiment involves inputting a current I to the ground and measuring the potential away from the source. In a time-domain system, the current has a duty cycle that alternates the direction of the current and has off-times between the current pulses at which the IP voltages are measured. A typical time-domain signature is shown in Figure 2. In that figure, ϕ_σ is the potential that is measured in the absence of chargeability effects. This is the "instantaneous" potential measured when the current is turned on. In mathematical terms

$$\phi_\sigma = \mathcal{F}_{dc}[\sigma], \tag{1}$$

where the forward mapping operator $\mathcal{F}_{dc}$ is defined by the equations

$$\nabla \cdot (\sigma \nabla \phi_\sigma) = -I\delta(\mathbf{r} - \mathbf{r}_s) \tag{2}$$
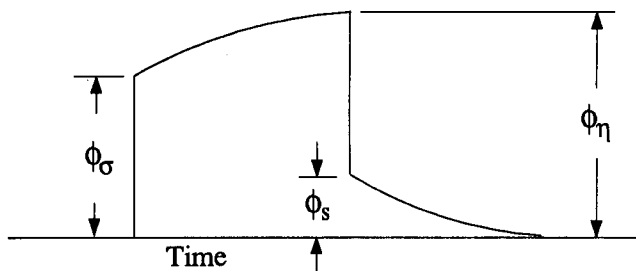

FIG. 2. Definition of the three potentials associated with the IP survey.

and the boundary conditions are $\partial \phi_\sigma / \partial n = 0$ at the earth's surface and $\phi_\sigma \to 0$ as $|r - r_s| \to \infty$, where $\mathbf{r}_s$ denotes the location of the current electrode.

If the ground is chargeable, then the potential ϕ_η, recorded after the constant current is applied, will differ from ϕ_σ. According to Siegel's formulation, the effect of the chargeability of the ground is modeled by using the DC resistivity forward mapping $\mathcal{F}_{dc}$ but with the conductivity replaced by $\sigma = \sigma(1 - \eta)$. Thus

$$\phi_\eta = \mathcal{F}_{dc}[\sigma(1 - \eta)], \tag{3}$$

or

$$\nabla \cdot (\sigma(1 - \eta)\nabla \phi_\eta) = -I\delta(\mathbf{r} - \mathbf{r}_s). \tag{4}$$

The IP datum, which we refer to as apparent chargeability, is defined by

$$\eta_a = \frac{\phi_s}{\phi_\eta} = \frac{\phi_\eta - \phi_\sigma}{\phi_\eta} \tag{5}$$

or

$$\eta_a = \frac{\mathcal{F}_{dc}[\sigma(1 - \eta)] - \mathcal{F}_{dc}[\sigma]}{\mathcal{F}_{dc}[\sigma(1 - \eta)]}. \tag{6}$$

Equation (6) shows that the apparent chargeability can be computed by carrying out two DC resistivity forward modelings with conductivities σ and $\sigma(1 - \eta)$.

Equation (6) defines the forward mapping for the IP data. The data can be inverted by: (1) linearizing this equation and solving a linear inverse problem, (2) introducing a formal DC inverse operator $\mathcal{F}_{dc}^{-1}$ and obtaining η by manipulating the conductivity models obtained after applying $\mathcal{F}_{dc}^{-1}$ to ϕ_η and to ϕ_σ, or (3) solving a nonlinear inverse problem that involves linearization but iterates until data predicted from equation (6) are in agreement with the observations. Irrespective of the method, the inversion of the IP data requires that the potentials ϕ_σ be inverted to recover the background conductivity. We address this issue first.

INVERSION OF DC POTENTIALS

Equation (1) is a nonlinear relationship between the observed potentials ϕ_σ and the conductivity σ. The goal of the inverse problem is to find the function σ which gave rise to those observations. There have been many attacks on this problem. For the example problem in this paper where the structure is presumed to be 2-D, we use the subspace inversion method given in Oldenburg et al. (1993). This methodology will be used for inverting both DC data and IP data, and we therefore outline essential details of the approach.

Let the data be denoted generically by the symbol d and the model by m. To carry out forward modeling to generate theoretical responses, and also to attack the inverse problem, we divide our model domain into M rectangular cells and assume that the conductivity is constant within each cell. Our inverse problem is solved by finding the vector $\mathbf{m} = \{m_1, m_2, \ldots, m_M\}$ which adequately reproduces the observations $\mathbf{d}_0 = (d_{01}, d_{02}, \ldots, d_{0N})$.

The inverse problem is posed as a standard optimization:

minimize $\psi_m(\mathbf{m}, \mathbf{m}_0) = \|\mathbf{W}_m(\mathbf{m} - \mathbf{m}_0)\|^2$

subject to $\psi_d(\mathbf{d}, \mathbf{d}_0) = \|\mathbf{W}_d(\mathbf{d} - \mathbf{d}_0)\|^2 = \psi_d^*$.   (7)

In equation (7), $\mathbf{m}_0$ is a base model and $\mathbf{W}_m$ is a general weighting matrix that is designed so that a model with specific characteristics is produced. The minimization of the model objective function ψ_m yields a model that is close to $\mathbf{m}_0$ with the metric defined by $\mathbf{W}_m$ and so the characteristics of the recovered model are directly controlled by these two quantities. The choice of ψ_m is crucial to the solution of the inverse problem, but we defer the details until later. $\mathbf{W}_d$ is a datum weighting matrix. We shall assume that the noise contaminating the jth observation is an uncorrelated Gaussian random variable having zero mean and standard deviation ε_j. As such, an appropriate form for the $N \times N$ matrix is $\mathbf{W}_d = \text{diag}\{1/\varepsilon_1, \ldots, 1/\varepsilon_N\}$. With this choice, ψ_d is the random variable distributed as chi-squared with N degrees of freedom. Its expected value is approximately equal to N and accordingly, ψ_d^*, the target misfit for the inversion, should be about this value. The appropriate objective function to be minimized is

$$\psi(\mathbf{m}) = \psi_m(\mathbf{m}, \mathbf{m}_0) + \mu(\psi_d(\mathbf{d}, \mathbf{d}_0) - \psi_d^*),   \quad (8)$$

where μ is a Lagrange multiplier.

The inverse problem is nonlinear and is generally attacked by linearizing equation (8) about the current model $\mathbf{m}^{(n)}$, differentiating with respect to parameters m_j and solving the resultant $M \times M$ system of equations for a perturbation $\delta\mathbf{m}$. This can be computationally intensive when M becomes large and hence the use of the subspace methods. In the subspace method, the "model" perturbation is restricted to be a linear combination of search vectors $\{\mathbf{v}_i\}$ $i = 1, q$. Thus

$$\delta\mathbf{m} = \sum_{i=1}^{q} \alpha_i \mathbf{v}_i \equiv \mathbf{V}\boldsymbol{\alpha}. \quad (9)$$

The linearized objective function, obtained by substituting $\mathbf{m}^{(n)} + \mathbf{V}\boldsymbol{\alpha}$ into equation (8), is

$$\psi(\mathbf{m}^{(n)} + \mathbf{V}\boldsymbol{\alpha}) = \psi_m + \boldsymbol{\gamma}_m^T \mathbf{V}\boldsymbol{\alpha} + \tfrac{1}{2}\boldsymbol{\alpha}^T\mathbf{V}^T\mathbf{J}^T\mathbf{J}\mathbf{V}\boldsymbol{\alpha} +$$

$$\mu\{\psi_d + \boldsymbol{\gamma}_d^T\mathbf{V}\boldsymbol{\alpha} + \tfrac{1}{2}\boldsymbol{\alpha}^T\mathbf{V}^T\mathbf{W}_d^T\mathbf{W}_d\mathbf{V}\boldsymbol{\alpha} - \psi_d^*\},$$

(10)

where $\boldsymbol{\gamma}_m = \nabla_m\psi_m$ and $\boldsymbol{\gamma}_d = \nabla_m\psi_d$ are gradient vectors, and ∇_m is the operator $(\partial/\partial m_1, \ldots, \partial/\partial m_M)^T$. In equation (10), ψ_m is understood to be $\psi_m(\mathbf{m}^{(n)}, \mathbf{m}_0)$, ψ_d is $\psi_d(\mathbf{d}^{(n)}, \mathbf{d}_0)$, and the sensitivity matrix $\mathbf{J}$ has elements $J_{ij} = \partial d_i(m^{(n)})/\partial m_j$. Differentiating equation (10) with respect to $\boldsymbol{\alpha}$ and setting the resultant equations equal to zero yields

$$\mathbf{V}^T(\mathbf{W}_m^T\mathbf{W}_m + \mu\mathbf{J}^T\mathbf{J})\mathbf{V}\boldsymbol{\alpha} = -\mu\mathbf{V}^T\boldsymbol{\gamma}_d - \mathbf{V}^T\boldsymbol{\gamma}_m. \quad (11)$$

The solution of this system requires that a line search be carried out to find the value of the Lagrange multiplier μ so that a specific target value ψ_d^* is achieved. This involves an initial guess for μ, solving equation (11) by SVD for the vector $\boldsymbol{\alpha}$, computing the perturbation $\delta\mathbf{m}$, carrying out a forward modeling to evaluate the true responses and misfit, and adjusting μ. The estimation of an acceptable value of μ typically requires three or four forward modelings.

The solution of equation (11) requires that $\mathbf{V}^T(\mathbf{W}_m^T\mathbf{W}_m + \mu\mathbf{J}^T\mathbf{J})\mathbf{V}$ be inverted and the numerical efficiency of the inversion is therefore realized since this is a $q \times q$ matrix. At each iteration in the inversion, we desire a model perturbation that minimizes ψ_m and alters ψ_d so that it achieves a specific target value ψ_d^*. To prevent the buildup of unnecessary roughness, the target misfit begins at an initial value (usually a fraction of the misfit generated by the starting model) and decreases with successive iterations towards a final value selected by the interpreter. Convergence is reached when the data misfit reaches this final target and no further reduction in the model norm is obtained with successive iterations.

In the IP survey carried out to produce Figure 1, surface electrodes are located every 10 m in the interval $x = (-100, 100\text{ m})$. Each of the 21 electrode positions can be activated as a current site and when it is, electric potentials are recorded at the remaining electrodes. The observed data set consists of 420 potential values, each of which has been contaminated by Gaussian noise having a standard deviation equal to 5 percent of the true potential. The data are generated using a finite-difference code (McGillivray, 1992), and the mesh, used both for forward modeling and for the inversion, consists of 1296 elements.

Because of the nonuniqueness of the inverse solution, the character of the final model is heavily influenced by the model objective function. Our choice for ψ_m is guided by the fact that we often wish to find a model that has minimum structure in the vertical and horizontal directions and at the same time is close to a base model m_0. This model, because it is "simple" in some respect, may well be representative of the major earth structure; however, other earth models might be closer to reality. Also, even if a geologically reasonable model has been found, it is insightful to generate different models that fit the data. This can provide understanding about whether features observed in the constructed model are required by the observations or if they are merely the result of minimizing a particular model objective function. An objective function that has the flexibility to accomplish these goals is

$$\psi_m(m, m_0) = \alpha_s \int\int w_s(m - m_0)^2 \, dx \, dz$$

$$+ \int\int \left\{ \alpha_x w_x \left(\frac{\partial(m - m_0)}{\partial_x}\right)^2 \right.$$

$$\left. + \alpha_z w_z \left(\frac{\partial(m - m_0)}{\partial_z}\right)^2 \right\} \, dx \, dz. \quad (12)$$

In equation (12), the functions w_s, w_x, w_z are specified by the user, and the constants α_s, α_x, α_z control the importance of closeness of the constructed model to the base model m_0 and control the roughness of the model in the horizontal and vertical directions. The discrete form of equation (12) is

$$\psi_m = (\mathbf{m} - \mathbf{m}_0)^T\{\alpha_s\mathbf{W}_s^T\mathbf{W}_s + \alpha_x\mathbf{W}_x^T\mathbf{W}_x + \alpha_z\mathbf{W}_z^T\mathbf{W}_z\}(\mathbf{m} - \mathbf{m}_0)$$

(13)

$$\equiv (\mathbf{m} - \mathbf{m}_0)^T\mathbf{W}_m^T\mathbf{W}_m(\mathbf{m} - \mathbf{m}_0),$$

and this defines the matrix $\mathbf{W}_m$ in equation (7). For the inversion of the synthetic data, we have set w_s, w_x, w_z equal to unity, and so in equation (13) $\mathbf{W}_s$ is a diagonal matrix with elements $\sqrt{\Delta x \Delta z}$, where Δx is the length of the cell and Δz is its thickness; $\mathbf{W}_x$ has elements $\pm\sqrt{\Delta z/\delta x}$, where δx is the distance between the centers of horizontally adjacent cells; and $\mathbf{W}_z$ has elements $\pm\sqrt{\Delta x/\delta z}$, where δz is the distance between the centers of vertically adjacent cells. In defining the objective function, we can choose m to be σ or $\ln \sigma$. Because earth conductivities typically vary over many orders of magnitude, we choose $m = \ln \sigma$. This choice also ensures that the recovered conductivity will remain positive. In addition, we specify $\alpha_s = .0002$, $\alpha_x = 1.0$, $\alpha_z = 1.0$ and the reference model m_0 corresponding to a background conductivity of $\sigma_0 = 5.0$ mS/m.

Our search vectors for the subspace inversion are obtained by partitioning ψ_d into data sets associated with individual current electrodes. Steepest descent vectors associated with these 21 data objective functions are combined with the steepest descent vector for the model gradient and a constant vector to form a basis for the subspace. The inversion begins with a halfspace of conductivity 5.0 mS/m. At every iteration we ask for a 50 percent decrease in the misfit objective function until a final misfit $\psi_d = N$ is achieved, where N is the number of data. A line search using forward modeling ensures that this is achieved, or in cases where it is not achievable, the line search is used to find that value of μ which provides the greatest decrease in the misfit. Once the target misfit has been obtained, the line search ensures that the misfit remains at the target value, and hence subsequent iterations alter only ψ_m. The desired misfit $\psi_d = 420$ is achieved by iteration 13, but a few more iterations are carried out until no further decrease in the model objective function is obtained. The model obtained at iteration 20 is shown in Figure 3c. It compares favorably with the true model in Figure 3a. The surface variation is well defined and so is the conductive anomaly in the center of the figure. There is no manifestation of the resistive ledge at the bottom left of the picture, but this might have been expected since its depth is 67 m and the electrodes span the region $(-100, 100)$ m.

INVERSION OF IP DATA

We shall develop three procedures for inverting IP data. Method I assumes that the chargeability $\eta \ll 1.0$. Equation (6) is linearized about the conductivity σ_b recovered from the inversion of the DC potentials, and a linear inverse problem is solved to recover η. In Method II we rearrange the mapping defined by equation (6) and compute η by manipulating the conductivities obtained by performing two DC resistivity inversions. In Method III we present a general approach that does not require that chargeability be small and solves the IP problem as a nonlinear inverse problem.

Method I: Linearization of the data equations

Let the earth model be partitioned into M cells and let η_i and σ_i denote the chargeability and electrical conductivity of the ith cell. Linearizing the potential ϕ_η about the conductivity model σ yields

$$\phi_\eta = \phi(\sigma - \eta\sigma) = \phi(\sigma) - \sum_{j=1}^{M} \frac{\partial \phi}{\partial \sigma_j} \eta_j \sigma_j + \text{H.O.T.} \quad (14)$$

Substituting into equation (6) yields

$$\eta_a = \frac{\phi_\eta - \phi_\sigma}{\phi_\eta} = \frac{-\sum_j \dfrac{\partial \phi}{\partial \sigma_j} \eta_j \sigma_j}{\phi(\sigma) - \sum_j \dfrac{\partial \phi}{\partial \sigma_j} \eta_j \sigma_j}. \quad (15)$$

This can be written approximately as

$$\eta_a = -\sum_j \frac{\sigma_j}{\phi} \frac{\partial \phi}{\partial \sigma_j} \eta_j = -\sum_j \frac{\partial \ln \phi}{\partial \ln \sigma_j} \eta_j. \quad (16)$$

Thus the ith datum is

$$\eta_{a_i} = \sum_{j=1}^{M} J_{ij}\eta_j, \quad i = 1, \ldots, N, \quad (17)$$
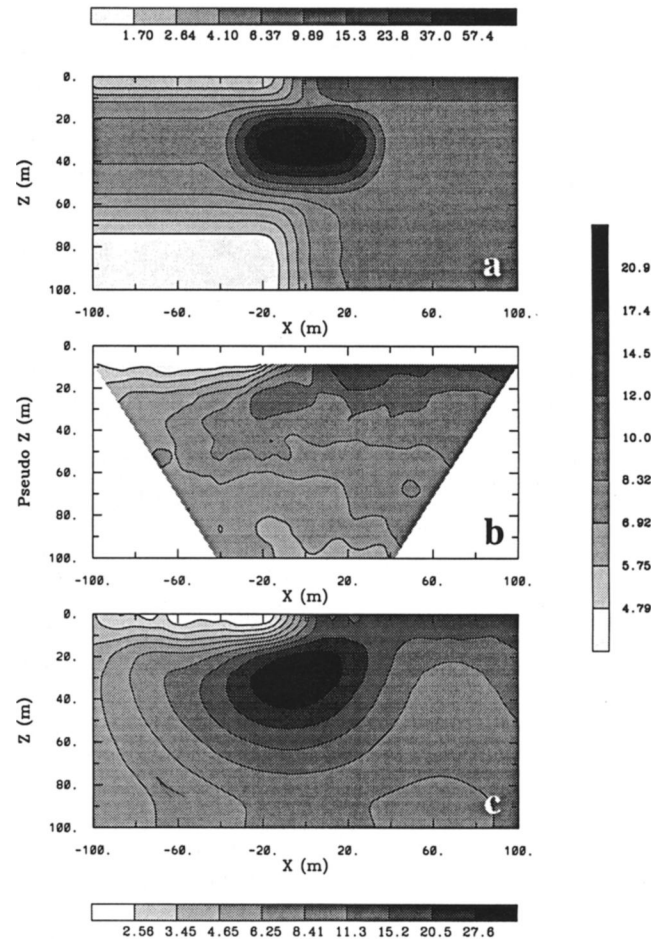

FIG. 3. The true and apparent conductivities are plotted respectively in (a) and (b). Inversion of the DC potentials yields the recovered model in (c). The grey scales indicate conductivity in mS/m.

where

$$J_{ij} = -\frac{\partial \ln \phi^i[\sigma]}{\partial \ln \sigma_j} \qquad (18)$$

is ijth element of the sensitivity matrix. In practice, the true conductivity σ is not known, and hence the background conductivity σ_b recovered in the DC resistivity inversion is used to compute the sensitivities in equation (18).

In the inversion of the DC potentials, we ensured positivity of the conductivity by working with $m = \ln \sigma$ in the model objective function. That was a natural choice since conductivity varies over many orders of magnitude, and it is relative change in conductivity that is often geologically meaningful rather than absolute change. Minimizing an objective function that penalizes variation of the logarithm of conductivity can therefore yield a geologically interpretable cross-section. Intrinsic chargeability, although positive, is confined to the region [0, 1). We are usually interested in the absolute deviation of chargeability from zero, and this range is rather small. Generally, we are not interested in the variation of chargeability in the range between zero and some small number (e.g., 0.01), and working with logarithms can put undue emphasis on these small values, especially if there is a smallest model component in the objective function, that is, if ψ_m has a component with $\alpha_s \neq 0$. The basic question is really whether the final model is more easily interpretable as a logarithmic quantity or as a linear quantity. Our belief is that linear chargeability is likely to be more meaningful and hence ψ_m should be evaluated with $m = \eta$. We adopt this throughout most of the paper; however, there is a degree of subjectivity in our choice, and we therefore present an example using $m = \ln \eta$.

If $m = \eta$, then invoking positivity requires that an extra mapping be introduced since variables in the subspace solution are intrinsically positive or negative. The inclusion of this mapping is simply made in the subspace algorithm and details are relegated to the Appendix.

As a test example, we consider the synthetic chargeability model shown in Figure 1c. It consists of a chargeable layer of $\eta = 0.05$ at the surface and a chargeable block of $\eta = 0.15$ at depth. The position of the chargeable block has been offset laterally from the position of the conductive block shown in Figure 1a. Offsetting the chargeable block from the conductive block is physically realistic and it will also be used to emphasize the ability of our inversion to work in such situations.

The 420 IP data collected in the survey are contaminated with Gaussian noise having a standard deviation equal to .002. This corresponds to about 5 percent of the average value of the apparent chargeabilities. After setting $m = \eta$ and writing $\mathbf{W}_m$ as $\mathbf{W}_\eta$, the inverse problem is solved by minimizing the objective function

$$\psi = \|\mathbf{W}_\eta(\eta - \eta_0)\|^2 + \mu(\|\mathbf{W}_d(\mathbf{J}\eta - \eta_a^{obs})\|^2 - \psi_d^*),$$

$$(19)$$

where ψ_d^* is a target misfit. In the following inversions, $\mathbf{W}_\eta$ is identical to the model weight matrix used for inverting the DC resistivity data, and the reference chargeability model η_0 has been set to zero. The inversion is carried out with the

same 23 subspace vectors used in the DC inversion. The resultant chargeability model, which has a desired misfit of 420 and was obtained after 18 iterations, is shown in Figure 4c. There has been good recovery of both the chargeable surface layer and the subsurface block.

The inversion of IP data is a multistep process and consequently there are numerous reasons why our inversion result is not identical to the true model. In particular, the discrepancies between Figures 4c and 4a can be caused by: (1) the availability of limited inaccurate data, (2) the choice of model objective function, (3) choice of subspace vectors or other aspects of the subspace inversion algorithm, (4) incorrect sensitivities J_{ij} resulting from the fact that the estimated conductivity differs from the true conductivity, (5) incorrect sensitivities resulting from numerical calculation, and (6) errors caused by the neglect of second-order terms in the linearized expansion that led to equation (17). It is possible to gain some insight as to which of these factors might be the most important. To do so, we perform three more inversions. The purpose of the first inversion is to evaluate the differences between the true and recovered model that are attributable to specifics of the inversion algorithm. This includes the effect of limited inaccurate data, the form of the objective function to be minimized, and all aspects of the subspace inversion. The test is performed by generating a new set of data $\eta_a = \mathbf{J}\eta_t$, where $\mathbf{J}$ is the sensitivity used in the inversion that produced the result in Figure 4c and η_t is the true chargeability. The same realization of error has been added to the data and the inversion has been carried out with the same parameters. The model is shown in Figure 4d. There are differences between Figures 4c and 4d, but they are slight. The amplitude of the buried anomalous chargeable body in Figure 4c is slightly higher than in Figure 4d and there is slightly more roughness in the surface zone. This may have been caused by the slight roughness of the recovered conductivity in Figure 3c. The similarity in the inversions indicates that the primary difference between the true and recovered models in Figures 4a and 4c is a result of the model objective function and the limited inaccurate data. The algorithm is designed to produce a chargeability that has minimal structure in the horizontal and vertical directions subject to fitting erroneous data. As a result, the buried chargeable body does not achieve amplitudes as high as those of the true body and its boundaries are smoothed.

The next two inversions provide some indication about the effect of inaccurate estimation of the background conductivity used to generate $\mathbf{J}$. We first compute $\mathbf{J}$ from the true conductivity in Figure 1a. The inversion result is given in Figure 4e. The result is similar to that in Figure 4c even though the sensitivities have been developed from two different conductivities. It would seem that for purposes of inverting IP data, the conductivity obtained from inverting the DC data is an adequate approximation to the true conductivity. Last, we develop the sensitivities for a constant conductivity half-space. The sensitivities are independent of the value of the half-space, and the results from the inversion are shown in Figure 4f. There are significant differences between the models in Figures 4c and 4f. Nevertheless, the approximate location of the buried chargeable body is quite evident, particularly in the region that has

lower conductivity, and there is no question that this image is more easily interpretable than the pseudosection in Figure 4b. This is encouraging in that it suggests that a relatively crude approximation may serve as an adequate background conductivity for the IP inversion. This is an area of future research.

As a last investigation, we carry out the inversion with $m = \ln \eta$. It is not possible to incorporate a reference model of $\eta_0 = 0$ and so the objective function is modified by setting $\alpha_s = 0$. Otherwise ψ_m is identical to that used in the previous chargeability inversions. The $\ln \eta$ model is shown in Figure 5a. The chargeable body is clearly observed but its depth is somewhat greater than the true depth and that shown in Figure 4c. The chargeable region on the surface is also visible. Plotting the results in Figure 5a on a linear scale produces Figure 5b. The chargeable body has been contracted to a circular feature whose central value is about 40 percent greater than the true chargeability. Nevertheless, the essential features of the true chargeability model are clearly visible. To be precise with comparisons, we have rerun the inversion in Figure 4c with an objective function in which $\alpha_s = 0$. The result is given in Figure 5c, and it differs imperceptably from the result in Figure 4c which used $\alpha_s = .0002$. Comparison of Figures 4a, 5a, and 5c shows that the center of the chargeable body in Figure 5c coincides more closely with that of the true model, but its value is about 40 percent less than the true chargeability. The contours in Figure 5c exhibit some elongation in the horizontal direction, indicating that the body might be wider than it is thick. At a first level of interpretation, however, all three models in Figure 5 provide consistent information about earth chargeability, and certainly any one is preferable to making an interpretation from the pseudosection. We shall continue to use $m = \eta$ and make our inversions and interpretations directly with linear chargeability.

Method II: An exact formalism

Our second method for inverting IP data follows directly from the definitions in equations (1) and (3) and the formal introduction of an inverse mapping operator $\mathcal{F}_{dc}^{-1}$. Applying this operator to equations (1) and (3) respectively yields
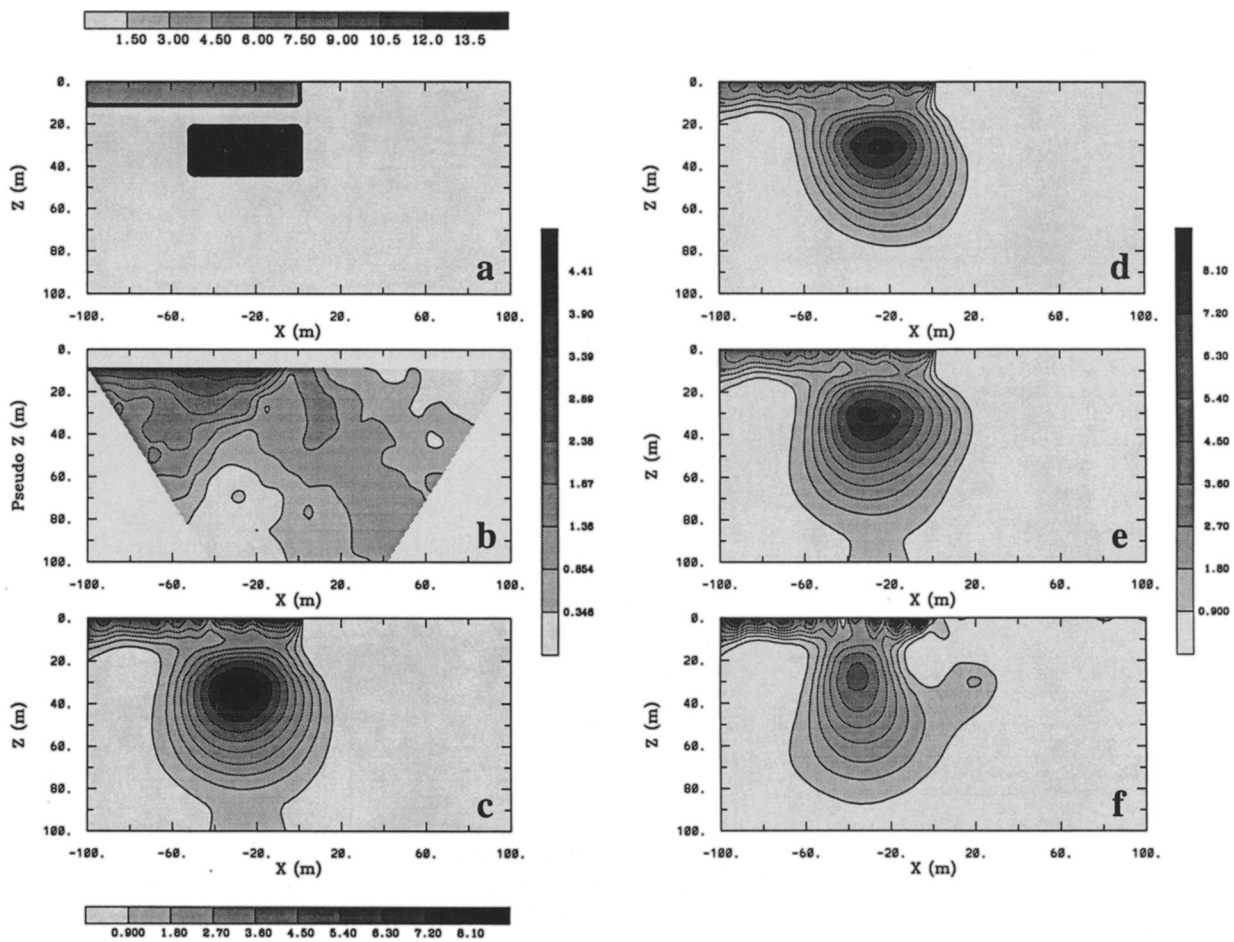

FIG. 4. Inversion results using Method I. The true and apparent chargeability models are shown in (a) and (b). An inversion that uses sensitivities obtained from the inverted conductivity in Figure 3(c) produces the chargeability model in (c). Discrepancies between (a) and (c) that are related to aspects of the subspace algorithm are inferred from (d). That panel shows the results of inverting synthetic data obtained by multiplying the sensitivity matrix $\mathbf{J}$ by the true chargeability model in (a) and inverting the resultant data. The chargeability models in (e) and (f) are the inversion results by respectively using the sensitivity matrix developed from the true conductivity model and the sensitivity matrix from a half-space.
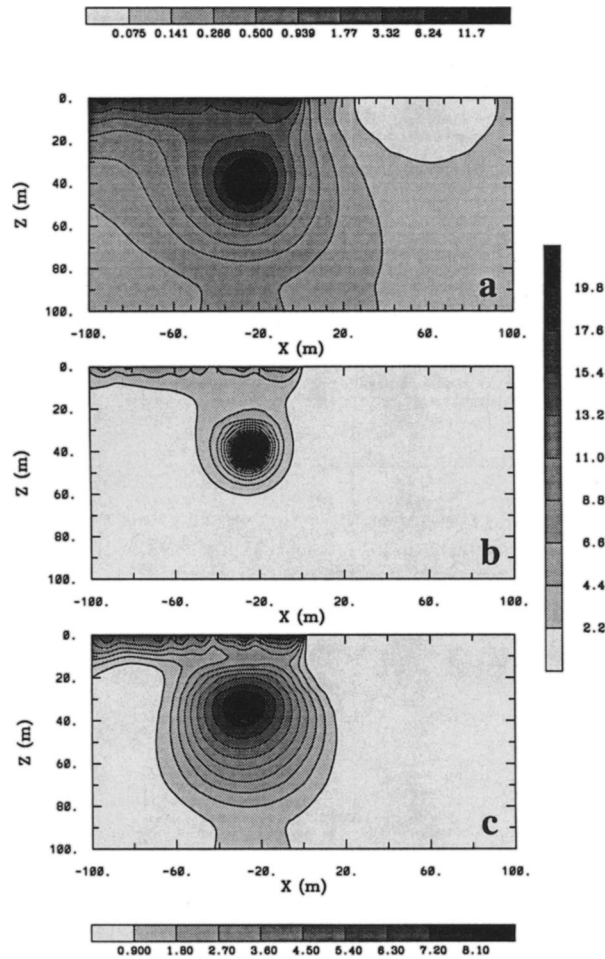
FIG. 5. Inversion results from Method I using $m = \ln \eta$ as the variable in the objective function. The recovered logarithmic chargeability is shown in (a) and it has been replotted in linear form in (b). For comparison, the inversion carried out using $m = \eta$ and setting $\alpha_s = 0$ is shown in (c). Thus (a) and (c) are a direct comparison of the results of using the same objective function form but applying it to $m = \ln \eta$ and to $m = \eta$, respectively.

$$\sigma = \mathcal{F}_{dc}^{-1}(\phi_\sigma), \qquad (20)$$

and

$$\sigma(1 - \eta) = \mathcal{F}_{dc}^{-1}(\phi_\eta), \qquad (21)$$

from which it follows that

$$\eta = \frac{\mathcal{F}_{dc}^{-1}(\phi_\sigma) - \mathcal{F}_{dc}^{-1}(\phi_\eta)}{\mathcal{F}_{dc}^{-1}(\phi_\sigma)}. \qquad (22)$$

Thus the chargeability model is produced by manipulating the conductivities obtained by inverting two sets of potential data taken in a DC/IP survey. To implement equation (22), we have inverted accurate ϕ_σ and ϕ_η data and have used the same model objective function that was used to recover the model in Figure 3c. It is important to minimize the same ψ_m; otherwise the nonuniqueness inherent in the inversion can allow radically different conductivities to be obtained and hence produce a poor estimate of η. The two inversions are

shown in Figures 6a and 6b. The differences between them are subtle and not readily visible to the eye. The difference between the conductivity models is shown in Figure 6c, and the chargeability recovered by applying equation (22) is given in Figure 6d. It is a rather good representation of the true chargeability.

Method III: Nonlinear inversion

The third method for inverting IP data is the most precise in that it solves a nonlinear inverse problem. With this approach, there is no necessity for the chargeability to be small. The inverse problem is stated as

minimize $\quad \psi_m(\eta, \eta_0) = \|\underline{W}_\eta(\eta - \eta_0)\|^2$

subject to $\quad \psi_d(\mathbf{d}, \mathbf{d}_0) = \|\underline{W}_d(\mathbf{d} - \mathbf{d}_0)\|^2 = \psi_d^*,$ $\qquad (23)$

where the data $\mathbf{d}$ are now the apparent chargeabilities, and the equation for producing the predicted data from any chargeability model η is given by equation (6) rather than equation (18) as it was with Method I.

The inversion proceeds by linearizing the data equations about the current model $\eta^{(n)}$ and computing the sensitivities $J_{ij} = \partial d_i / \partial \eta_j$. As in all of our methods, it is understood that the background conductivity model has been recovered by a previous inversion of the DC resistivity data. Writing the data equations as

$$d_i = \frac{\phi_\eta^i - \phi_\sigma^i}{\phi_\eta^i} \qquad (24)$$

and differentiating with respect to the parameter η_j yields

$$\frac{\partial d_i}{\partial \eta_j} = \frac{\phi_\sigma^i}{(\phi_\eta^i)^2} \frac{\partial \phi_\eta^i}{\partial \eta_j}. \qquad (25)$$

We therefore need only evaluate $\partial \phi_\eta^i / \partial \eta_j$. Writing $\sigma = \sigma_b(1 - \eta)$ as the conductivity that produces the potential ϕ_η, then

$$\frac{\partial \phi_\eta^i}{\partial \eta_j} = -\sigma_{bj} \frac{\partial \phi_\eta^i}{\partial \sigma_j} \equiv -\sigma_{bj} G_{ij} \qquad (26)$$

is simply a scaled value of the sensitivity for a DC resistivity problem. The final sensitivity is

$$J_{ij} = \frac{\partial d_i}{\partial \eta_j} = -\sigma_{bj} \frac{\phi_\sigma^i}{(\phi_\eta^i)^2} G_{ij}. \qquad (27)$$

To solve the inverse problem, we appeal to the subspace methodology and carry out an identical procedure used to invert the potentials ϕ_σ (see section: Inversion of DC Potentials). In fact, because of the relationship between the IP sensitivities and the sensitivities for the DC problem given by equation (27), the algorithm to invert DC data requires only minor changes so that IP data can be inverted. When inverting the IP data we use 23 search vectors: 21 steepest descent vectors associated with data groupings from each current electrode, the steepest descent vector associated with the model objective function, and a constant vector. We use the same model objective function that produced the results in Figure 4. The positivity constraint in the subspace formulation has been implemented by the transformation in

the Appendix. The result of inverting the test data set, achieved after 23 iterations, is shown in Figure 7b. There has been a good recovery of the chargeability model. A comparison between Figures 4c and 7b quantifies the difference between solving a linearized problem and solving the complete nonlinear inverse problem for this example. The differences are small but this might be expected since the maximum chargeability in the model is only 0.15.

NONUNIQUENESS AND HYPOTHESIS TESTING IN IP INVERSIONS

Minimization of a particular objective function provides a single model from which to make geologic inferences. Without further investigation it cannot be determined which features of the constructed model are demanded by the observations and which are the result of the objective function that has been minimized. Some insight can be achieved by altering the objective function and carrying out other inversions. We present two examples that illustrate the potential usefulness of the flexible model objective function given in equation (12).

In some geologic environments it might be expected that the earth has greater continuity in the lateral direction compared to the vertical direction. To see the effect of this on the current example we carry out the inversion with $\alpha_x = 1.0$ and $\alpha_z = 0.1$. The result is shown in Figure 8a. Comparison with Figure 4c shows that the surface structure is smoothed in the horizontal direction and that the buried anomaly has been elongated. The transition from the bottom of the buried anomaly to the background is more rapid and the depth extent of the anomaly is reduced. The chargeabil-

ity low between the surface layer and the buried prism is better defined. Redoing the inversion with $\alpha_z = 0.01$, as shown in Figure 8b, further enhances these effects. It is noted that the horizontal extent and thickness of the surface chargeable layer, and the central position of the buried anomaly, remain unchanged from those recovered in the previous inversions. These results suggest that in this example, the locations of the recovered anomalies are primarily controlled by the data and not altered by different objective functions. This is further illustrated by the next example.

In the synthetic example, the chargeable prism has been offset from the conductive prism. The inverted chargeabilities reflected this shift. For some geologic settings, however, it might be expected that locations of high conductivity and chargeability coincide. To see if this hypothesis is compatible with the data example analyzed here, we have rerun the inversion but have introduced a weighting function that attempts to make the chargeability coincide with the conductive prism. This is done by introducing weighting functions, derived from the recovered conductivity anomaly, that attempt to force the recovered chargeability anomaly to coincide with those regions that differ from the background conductivity. To form the weight functions, we first convert the conductivity model σ_b to $r_\sigma = |\log_{10}(\sigma_b/\sigma_0)|$, where $\sigma_0 = 0.01$ S/m for this example. This quantity is then scaled to the range of $(0.02, 1)$ and its reciprocal is generated to form w_s. The weighting functions w_x and w_z for the derivative terms are generated in a similar manner from the gradient of r_σ so that the low values for w_x and w_z correspond to regions of high gradient in the recovered conductivity model. As an example, the function w_s gener-
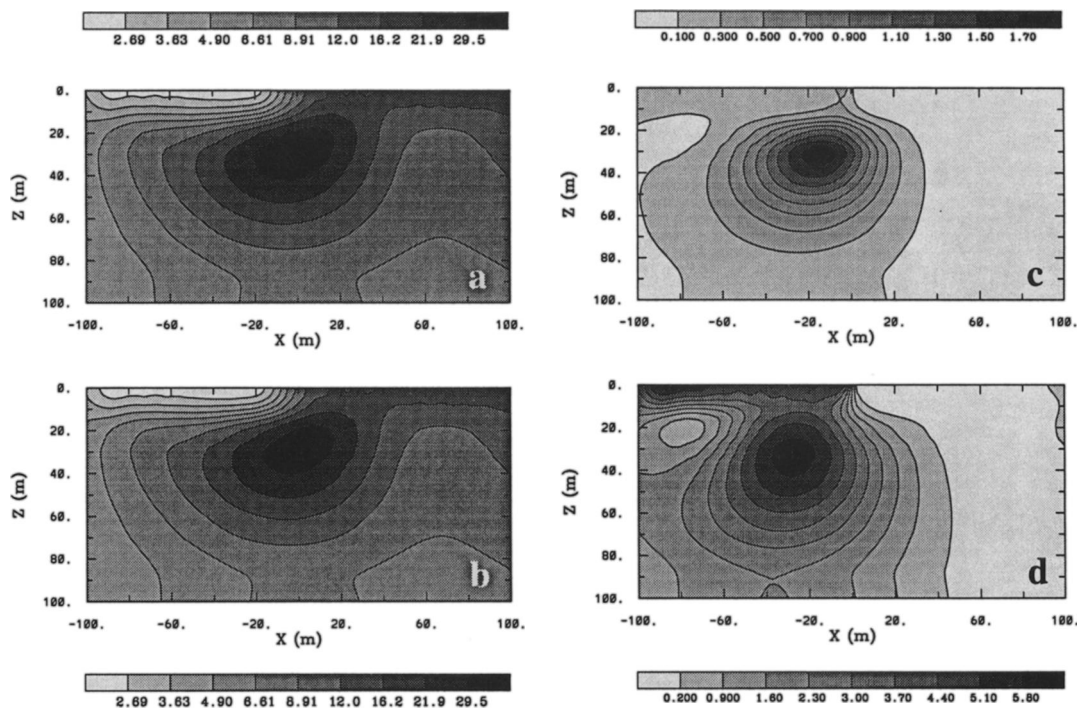

FIG. 6. Inversion results from Method II: The recovered conductivity structures obtained by inverting ϕ_σ and ϕ_η are shown in (a) and (b), respectively. The difference is shown in (c), and the recovered chargeability is shown in (d). The gray scales indicate conductivity in mS/m and chargeability in percent.

ated from the conductivity model in Figure 3c is shown in Figure 9a. The function has two zones of reduced weights corresponding to the surface resistive zone and the buried conductive prism. The latter would attempt to center the buried chargeability anomaly at $x = 0$. The model obtained by carrying out the weighted inversion is shown in Figure 9b and this can be compared with the unweighted inversion replotted in Figure 9c. The attempt to make the chargeability high coincide with the conductive high has not been successful. There has been some distortion of the contours toward the conductor, but the center of chargeability high remains close to its original and true position of $x = -25$ m. This provides greater confidence that the location of the chargeable high is demanded by the data and is not an artifact of the objective function that is being minimized.

PRACTICAL CONSIDERATIONS WITH RESPECT TO FIELD DATA

In developing the IP inversion algorithms we have made the assumption that the IP response is the apparent chargeability obtained from an idealized time domain experiment. Using Figure 2, the apparent chargeability is defined as the ratio of the secondary potential ϕ_s to the total potential measured ϕ_η that is measured just before the current is cut off. This is a dimensionless number and usually has a range $(0, 0.5)$. In reality, ϕ_s and ϕ_η are not recorded directly with the field acquisition system. In time-domain systems, ϕ_η is measured and then either a single sampling of $\phi_s(t)$ at some time beyond the current shutoff, or an integration of $\phi_s(t)$ over a time window, is recorded and used as an IP datum. In

a frequency-domain system, sinusoidal current waveforms are input to the ground at a low-frequency ω_ℓ and at a high frequency ω_h, and a percent frequency effect (PFE) is used for the IP response. In a phase IP system, the phase lag between a transmitted sinusoidal current and a received voltage can be measured directly. Although each of these systems measures something related to IP potentials, they do not measure the same quantity nor do any of them measure the quantity that is required for our inversions. In fact, the various data, still generally called "apparent chargeability," can be dimensionless or have units of ms or mrad.

To use the algorithms presented here, it is necessary to convert collected IP data into the form defined by equation (5). The only exception is when relative chargeabilities are required as output; that is, if the final distribution of chargeabilities can be in error by a constant factor. In such cases, Method I, which solves a linear problem, can still be used. With that algorithm, if IP data have been scaled by a constant factor F, then the recovered model will also be scaled by F and the output model has the same units as the data.

If the output chargeability model is to be calibrated, or if Methods II or III are used, then the data collected from any IP experiment must be converted to the response used here. This can be done. A study of the relationships between these parameters is given in Van Voorhis et al. (1973) and Nelson and Drake (1973).

Another practical aspect about data collection is that the potential ϕ_σ is not generally measured. Yet it is this potential that should be inverted to recover the background conductivity. Usually the potential ϕ_η is measured and inverted to recover the background conductivity, but the inversion of
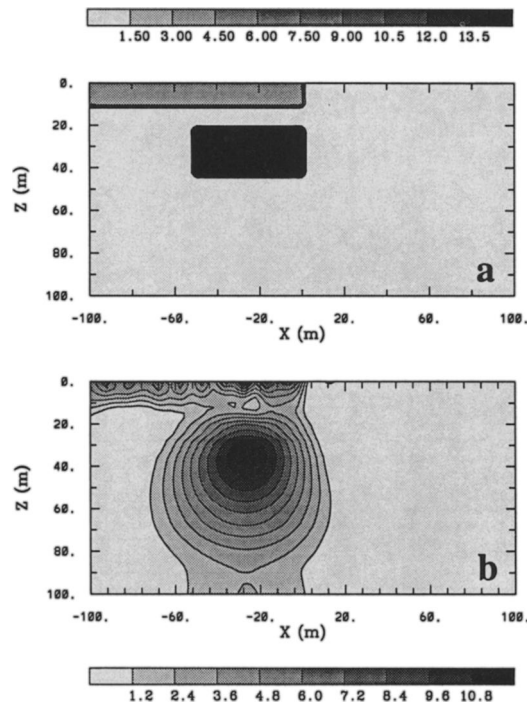

FIG. 7. Inversion results from Method III. The true and recovered chargeabilities are shown in (a) and (b), respectively.
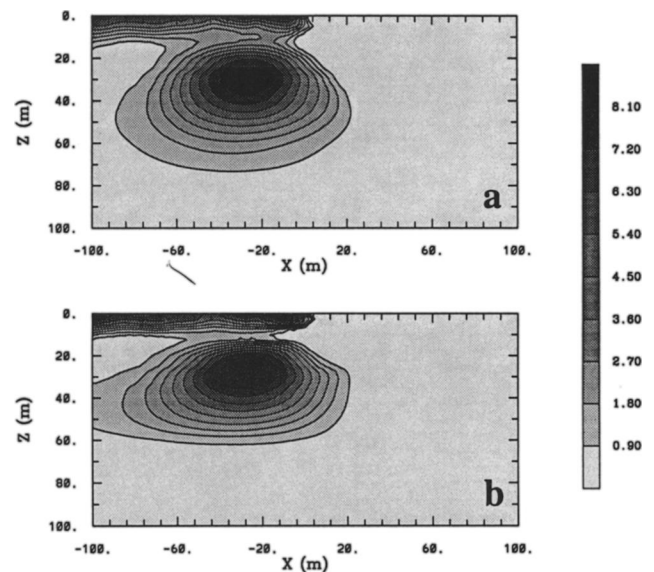

FIG. 8. Exploring nonuniqueness in IP inversions by altering the objective function. The model in (a) was obtained by setting $\alpha_s = 0.0$, $\alpha_x = 1.0$, and $\alpha_z = 0.1$. This has an increased penalty against variation in the horizontal direction and hence the chargeability is elongated horizontally. The effect is increased in (b) where $\alpha_z = 0.01$.

ϕ_η yields the conductivity $\sigma(1 - \eta)$. If η is small, then this difference may not be important; however, if η is large, then a correction may be demanded. Two possible routes exist. The first is to estimate the potentials ϕ_σ by computing $\phi_\eta - \phi_s$ from the field measurements. Although this procedure may enhance the error on the data, it does produce a datum that is ideally unaffected by the polarization of the subsurface. The second method is an iterative correction that could be performed one or more times and consists of steps: (1) perform a DC inversion on ϕ_η to yield $\sigma_{b\eta}$, (2) use $\sigma_{b\eta}$ to carry out the IP inversion to recover η, (3) form a new background conductivity $\sigma_b = \sigma_{b\eta}/(1 - \eta)$, and (4) solve the IP inverse problem again. We have not implemented the above correction procedures on our field data example presented in the next section because the chargeabilities are quite small. We do however, recognize its potential importance for some field data.

A matter of practical importance with respect to any of our inversion methods is the issue of ascribing errors to the data and deciding upon the degree to which the data are to be fit.
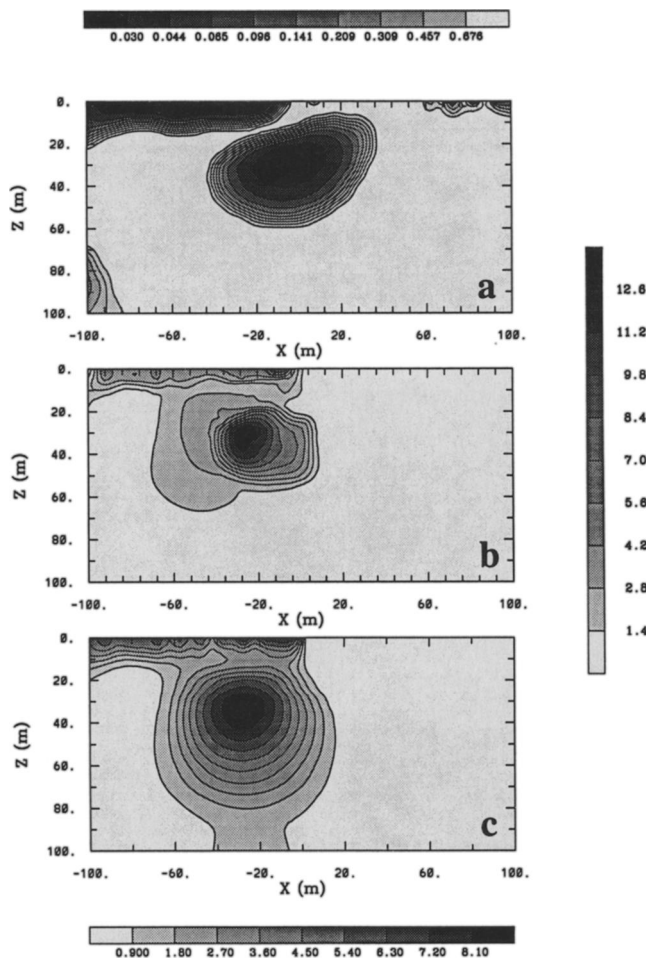

FIG. 9. Hypothesis testing in IP inversions. A weighing function that attempts to move the chargeable zone so that it coincides with the high conductivity has been incorporated into the inversion. The weighting function is shown in (a) and the recovered model is given in (b). That can be compared with the model in (c) which has no preferential weighting.

We have assumed that data errors are Gaussian and independent. This is surely violated in practice. Also, the "errors" ascribed to the data must reflect not only measurement errors (additive noise, incomplete removal of electromagnetic coupling, electrode effects, etc.) but also geologic "noise" caused by 3-D bodies and mathematical cellularizations which are too coarse to emulate true geologic variability even if the earth is 2-D. As a consequence, it is not possible to assign precise standard deviations for the observations. Our approach is pragmatic. For the first inversion, we assign to each datum a standard deviation that is a percentage of the average apparent chargeability. If a model can be found that is in acceptable agreement with these assigned errors, then we adopt this model and carry out another inversion in which we assign somewhat lower standard deviations to see if more structure can be extracted from the data. Alternatively, if the inversion fails to find a model that adequately reproduces the observations, then we increase the assigned errors. Contoured maps of misfits between the observed and predicted data are invaluable in assessing appropriateness of assigned standard deviations and in detecting outliers. In summary, the question of error assignment is problematic and dictates that the inversion will be carried out a number of times.

FIELD DATA EXAMPLE

The inversions presented here have concentrated upon using pole-pole data. The techniques can, without alteration, be used to invert IP data from any electrode geometry. The field example shown next inverts dipole-dipole data from a gold region in the Abitibi clay belt in Quebec. The geologic target is a shear-zone-controlled gold mineralization associated with silicification and sulphides. The geophysical interpretation is complicated by the presence of a conductive overburden of variable thickness. The overburden may contain chargeable clays that can dilute or even mask completely the IP response from the subcrop. It is also likely that the effect upon the resistivity caused by silicification in the subcrop will be confused with the variable thickness of the conductive overburden.

It was felt that the geologic structure might be reasonably 2-D and hence amenable to inversion with our algorithms. A dipole-dipole survey with 25 m spacings and $n = 1$ to $n = 5$ was carried out. Figures 10a and 11a show the original pseudosections of apparent conductivity and apparent phase. The main feature in the apparent conductivity section is the change in conductivity as one proceeds from the resistive region to the south into the conducting overburden towards the north. Considering that the apparent conductivity pseudosection is rather smooth, we assume a 5 percent error in each of the 140 potential data. The 2-D earth structure is divided into 48×27 cells, and so the inverse problem is to recover the unknown conductivity for each of the 1296 cells. The reference model σ_0 is chosen to be a uniform half-space of 1 mS/m. This is based upon the prior knowledge about the deep geological structure from Placer Dome, which indicates resistive country rock. In developing the objective function [equation (12)], we set the functions w_s and w_z to unity throughout the model but choose a special weighting $w_x = (\Delta z/\delta x)^2$ near the surface. This

tends to suppress abrupt horizontal changes that might result from the elongated cells near the surface. The inverted conductivity, achieved after 20 iterations, is shown in Figure 10b. Several features are observed in the model. The major conductive overburden begins at 1525, extends northward, and attains a thickness of over 40 m. There is also a conductive layer south of 1525, but its thickness is much smaller and variable. The model shows two distinct resistive regions whose resistivities are almost twice that of the general background. The first is centered at 1525 and the second at 1300.

The IP pseudosection in Figure 11a displays considerable structure. There is a well-defined anomaly near 1825, and a larger region from 1300 to 1600 that contains high and variable chargeabilities. The data provided are apparent phases in milliradians. The IP inversion requires apparent chargeabilities (as per Siegel's definition) as data and inverts for the intrinsic chargeability model. Since to first order apparent phase is proportional to apparent chargeability, we have chosen to use apparent phases as data for the inversion. We use Method I for the inversion and hence the recovered model will also have units of mrad. The apparent phase data appear to be more noisy than the apparent resistivities, and this has made the assignment of an appropriate error more difficult. We have assigned a constant error equal to 0.13 mrad to each datum. This corresponds to 5 percent of the middle value of the range of the apparent phase. This "estimate" of the error is used in the inversion but, as we shall illustrate, it is likely too small.

The chargeable earth model is discretized into 48 by 27 cells so that 1296 unknown intrinsic chargeabilities are estimated from 140 data. The model objective function is identical to that in the DC inversion except the reference model η_0 is taken as zero. The target misfit is 140, but the subspace inversion code was not able to find a model that reproduced the data to that degree. The model shown in Figure 11d has the smallest misfit obtainable by the algorithm. The misfit is $\chi^2 = 1330$, and this translates into an average misfit of 0.4 mrad rather than 0.13 mrad. There is always uncertainty about the data errors and equivalently, the degree to which the data should be fit. In general, the features observed in the inversion will decrease in amplitude as the misfit is allowed to increase. To obtain some confidence about the features observed, we have carried out two more inversions. The respective misfits were 1400 and 1500 (i.e., 0.41, and 0.42 mrad), and the models are shown in Figures 11c and 11b. The anomalies become shallower and have reduced amplitude as the misfit is increased, but the three major IP bodies are evident in all inversions and this provides confidence that the bodies are not artifacts of overfitting the data.

Close inspection of Figures 10b and 11d shows that the IP anomalies are offset from the conductivity anomalies. We emphasize that these results illustrate the need for an algorithm that does not coarsely parameterize the models, but rather allows many cells and finds a solution by minimizing an objective function of the model.
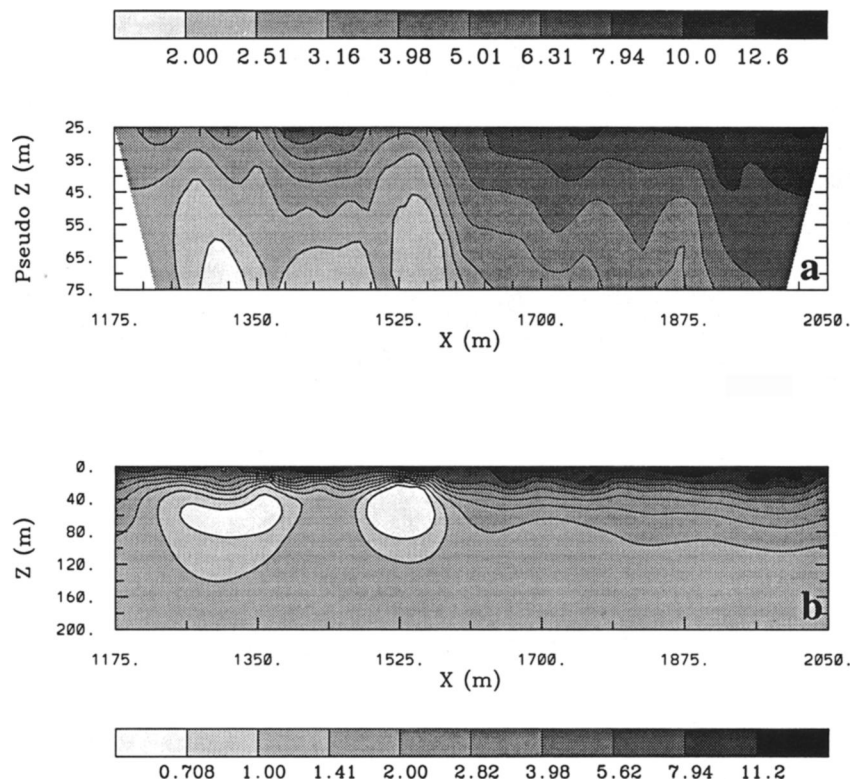

FIG. 10. A dipole-dipole apparent conductivity pseudosection is shown in (a). The inverted conductivity is presented in (b). Gray scales indicate the conductivity in mS/m.

DISCUSSION

We have presented three methods by which IP data can be inverted. The methods have been illustrated in two dimensions by inverting pole-pole and dipole-dipole data, but no alteration is required to invert any configuration of electrode data over 1-D or 3-D structures. All three methods require a prior inversion of DC potentials to recover a background conductivity σ_b. The following summarizes our thoughts about the three methods.

Method I.— The equations are linearized about the background conductivity. An advantage of this approach is that only a linear inverse problem needs to be solved and positivity is easily incorporated. The inversion is rapid and can be carried out many times with different model objective functions or with different estimates of data error and target misfits. Additionally, the inversion results are insensitive to scale errors of the data. These are very positive attributes. There is, however, an explicit need to compute and store the sensitivity matrix $\mathbf{J}$. Also, nonlinear effects are unaccounted for, and these can become important when chargeabilities are large.

Method II.— The appeal of this method is that only one computing algorithm to invert DC potentials is required. A
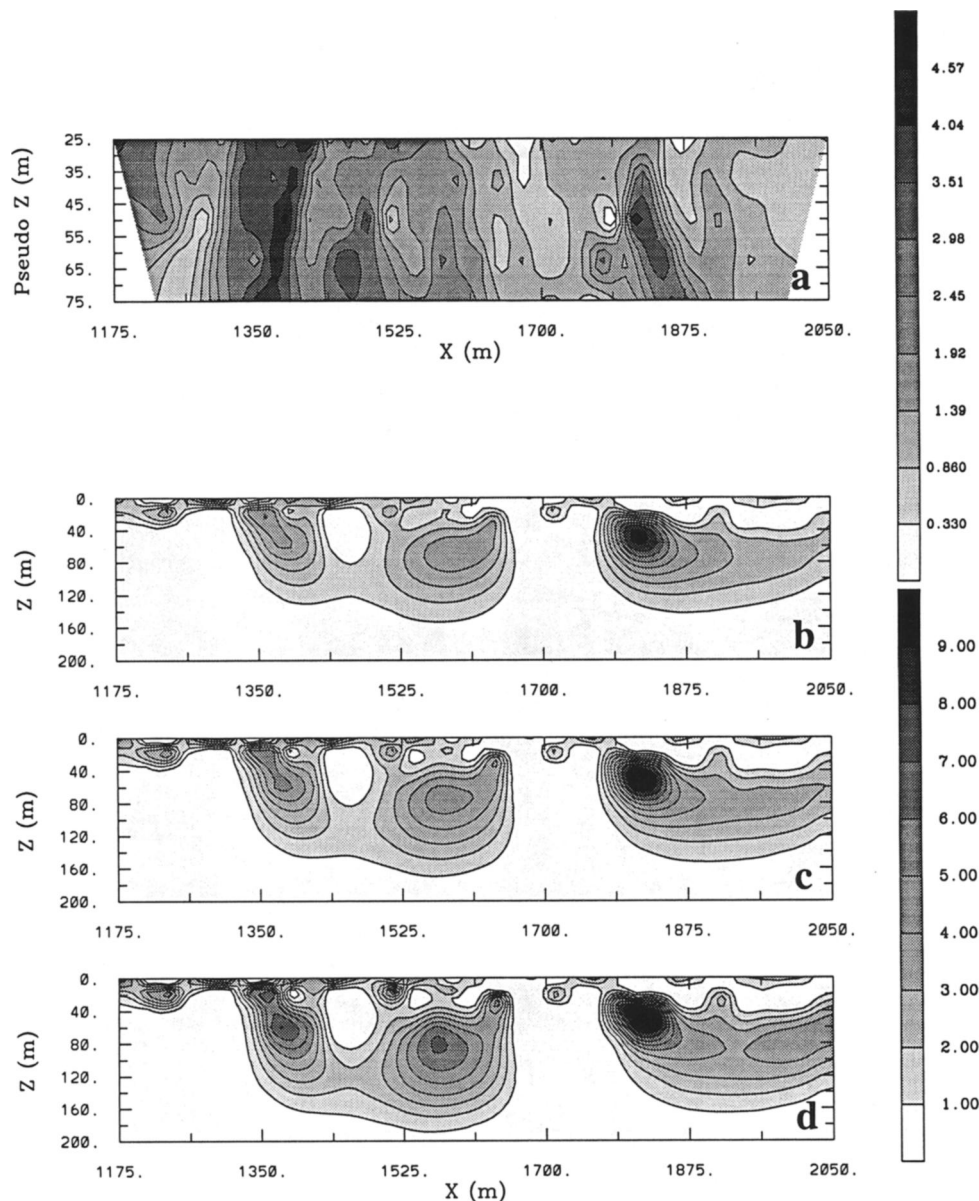

FIG. 11. The dipole-dipole apparent chargeability phase pseudosection is given in (a). Values are in mrad. Three inversions, carried out at successively smaller misfit levels are given in (b)–(d). Chi-squared misfits, assuming an initial standard deviation of each datum of 0.13 mrad, are, respectively, 1500, 1400, and 1330 for the three inversions.

difficulty, however, is a potential lack of robustness arising because the desired chargeability is a small quantity that is obtained by subtracting two large quantities. This requires that the inversion algorithm is stable, that is, $\mathcal{F}_{dc}^{-1}$ must be such that small changes in the data result in small changes in the conductivity model. The DC inversion algorithm used here is robust in this regard, but we still have had difficulty using this procedure with error contaminated data. Another difficulty with this method is that the two data sets ϕ_σ and ϕ_η are not generally recorded directly. Rather, ϕ_η and a number related to ϕ_s are acquired in field surveys.

Method III.— This has the advantage that it can handle large chargeabilities and that essentially the same algorithm can be used to invert both the DC and the IP data. This is especially true if positivity is implemented by using ln η as the variable in the inversion code. On the other hand, field observations must be converted to chargeabilities defined by ϕ_s/ϕ_η, and the inversion results are sensitive to any errors resulting from that transformation. Also, a complete nonlinear inversion must be carried out to compute the chargeability. This is a disadvantage compared to Method I if many inversions are to be completed with different model objective functions and/or solved according to different misfit levels.

We have no absolute preference for methodology to invert IP data. Both Methods I and III are viable. From a theoretical viewpoint, Method III is likely the best. However, for most of our inversions, we have used Method I, principally because it allows us to bypass the step of converting field data to true responses and because we can easily carry out a number of inversions at different misfit levels and with different objective functions. A similar lack of conclusiveness exists with the choice of whether η or ln η is used as the variable in the model objective function. For reasons already stated in the text, we prefer to use η, but images of ln η may be just as geologically interpretable.

The important aspect of our inversion methodology is that the generic model objective function that is minimized provides great flexibility to generate different models. With a properly designed objective function it is possible to incorporate additional information about the distribution of conductivity or chargeability and to generate a model that is in accordance with geologic constraints. Such a model may be regarded as a best estimate for the true earth structure and can be used in a final interpretation. However, altering the objective function and carrying out additional inversions allows exploration of model space and provides an indication of which features are demanded by the data. These two aspects of constructing a most-likely model, and exploring the range of acceptable models, form the foundation of a responsible interpretation.

ACKNOWLEDGMENTS

REFERENCES

Barker, R. D., 1990, Investigation of groundwater salinity by geophysical methods, *in* Ward, S. H., Ed., Geotechnical and environmental geophysics: Investigations in Geophysics No. 5, Vol 2, Soc. Expl. Geophys., 201–211.

Beard, L. P., and Hohmann, G. W., 1992, Subsurface imaging using approximate IP inversion: 62nd Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 427–430.

Bertin, J., and Loeb, J., 1976, Experimental and theoretical aspects of induced polarization, Vol. 1 and 2: Gebruder Borntraege.

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

Hohmann, G. W., 1990, Three-dimensional IP models, *in* Fink, J. B., McAlister, E. O., Sternberg, B. K., Wieduwilt, W. G., and Ward, S. H., Eds., Induced polarization: Applications and case histories: Investigations in geophysics No. 4, Soc. Expl. Geophys., 150–178.

Fink, J. B., McAlister, E. O., Sternberg, B. K., Wieduwilt, W. G., and Ward, S. H., (1990), Eds., Induced polarization: Applications and case histories: Investigations in Geophysics No. 4, Soc. Expl. Geophys.

La Brecque, D. J., 1991, IP tomography: 61st Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 413–416.

McGillivray, P. R., 1992, Forward modeling and inversion of DC resistivity and MMR data. Ph.D. thesis: University of British Columbia.

Nelson, P. H., and Drake, T. L., 1973, Letter to the editor regarding the paper "Complex resistivity spectra of porphyry copper mineralization": Geophysics, **38**, 984.

Oldenburg, D. W., and Li, Y., 1994, Subspace linear inverse method: Inverse Problems, **10**, 1–21.

Oldenburg, D. W., McGillivray, P. R., and Ellis, R. G., 1993, Generalized subspace methods for large scale inverse problems: Geophys. J. Int., **114**, 12–20.

Pelton, W. H., Rijo, L., and Swift, C. M., 1978, Inversion of two-dimensional resistivity and induced polarization data: Geophysics, **43**, 788–803.

Rijo, L., 1984, Inversion of 3-D resistivity and induced polarization data: 54th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 113–117.

Sasaki, Y., 1982, Automatic interpretation of induced polarization data over two-dimensional structures: Mem. Fac. Eng., Kyushu Univ., **42**, No. 1, 59–74.

Siegel, H. O., 1959, Mathematical formulation and type curves for induced polarization: Geophysics, **24**, 547–565.

Sumner, J. S., 1976, Principles of induced polarization for geophysical exploration: Elsevier Science Publ. Co., Inc.

Van Voorhis, G. D., Nelson, P. H., and Drake, T. L., 1973, Complex resistivity spectra of porphyry copper mineralization: Geophysics, **38**, 49–60.

Ward, S. H., Ed., 1990, Geotechnical and environmental geophysics: Investigations in Geophysics No. 5, Vol. 1–3, Soc. Expl. Geophys.

APPENDIX

INVOKING POSITIVITY IN INVERSE PROBLEMS

Positivity in subspace solutions is effected in the following manner. The reader is referred to Oldenburg and Li (1994) for more details. In an attempt to keep a uniform notation, let the symbol m denote the model for the inverse problem, so $m_i = \eta_i$ in what follows. To invoke positivity, let $m_i = f(p_i)$ and $p_i = f^{-1}(m_i)$ be mappings connecting the chargeability with new parameters p_i. Let $\mathbf{p}^{(n)}$ denote the model at the nth iteration. The value of the model objective function ψ_m evaluated at an updated model $\mathbf{p}^{(n)} + \delta\mathbf{p}$ is

$$\psi_m(\mathbf{p}^{(n)} + \delta\mathbf{p}) = \|\underline{\mathbf{W}}_m[f(\mathbf{p}^{(n)} + \delta\mathbf{p})] - \mathbf{m}_0\|^2. \tag{A-1}$$

By performing a Taylor expansion on f at the point $\mathbf{p}^{(n)}$ then

$$\psi_m = \|\underline{\mathbf{W}}_m\underline{\mathbf{F}}\delta\mathbf{p} + \underline{\mathbf{W}}_m(\mathbf{m}^{(n)} - \mathbf{m}_0)\|^2, \tag{A-2}$$

where $\mathbf{F}$ is a diagonal matrix with elements

$$F_{ii} = \frac{\partial f_i}{\partial p_i}\Big|_{p^{(n)}} = \frac{\partial m_i}{\partial p_i}\Big|_{p^{(n)}}. \tag{A-3}$$

A similar Taylor expansion applied to the misfit objective functional yields

$$\psi_d = \|\underline{\mathbf{W}}_d\underline{\mathbf{J}}\underline{\mathbf{F}}\delta\mathbf{p} + \underline{\mathbf{W}}_d(\mathbf{d}(\mathbf{p}^{(n)}) - \mathbf{d}^{obs})\|^2. \tag{A-4}$$

The minimization of equation (A-4) subject to $\psi_d = \psi_d^*$ is solved using the subspace formulation. The perturbation $\delta\mathbf{p}$ is represented as $\underline{\mathbf{V}}\alpha$ and the objective function $\psi = \psi_m + \mu(\psi_d - \psi_d^*)$ is solved by minimizing with respect to the coefficients α. This produces equations that are the same as those in equation (11) with the exception that the matrix $\underline{\mathbf{V}}$ has been replaced by the matrix $\underline{\mathbf{V}}\underline{\mathbf{F}}$.

The mapping used here is a two-segment mapping consisting of an exponential and a linear region

$$m = \begin{cases} 0, & p < p_b, \\ e^p - m_b, & p_b \le p \le p_1, \\ (p - p_1 + 1)e^{p_1} - m_b, & p > p_1, \end{cases} \tag{A-5}$$

where $p = p_1$ is the transition point between the two segments, and m_b is selected to be small enough such that model values smaller than m_b are not significantly different from zero.