

Three-dimensional magnetotelluric inversion: data-space method

Weerachai Siripunvaraporn^{a,*}, Gary Egbert^b, Yongwimon Lenbury^c,
Makoto Uyeshima^d

^a Department of Physics, Faculty of Science, Mahidol University, Rama VI Road, Rachatawee, Bangkok 10400, Thailand

^b College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331, USA

^c Department of Mathematics, Faculty of Science, Mahidol University, Rama VI Road, Rachatawee, Bangkok 10400, Thailand

^d Earthquake Research Institute, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan

Received 28 January 2003; received in revised form 24 October 2003; accepted 16 August 2004

Abstract

A three-dimensional magnetotelluric (MT) minimum structure inversion algorithm has been developed based on a data-space variant of the Occam approach. Computational costs associated with construction and inversion of model-space matrices make a model-space Occam approach to 3D MT inversion impractical. These difficulties are overcome with a data-space approach, where matrix dimensions depend on the size of the data set, rather than the number of model parameters. With the transformation to data space it becomes feasible to invert modest 3D MT data sets on a PC. To reduce computational time, a relaxed convergence criterion is used for the iterative forward modeling code used to compute the sensitivity matrix. This allows reduction in computational time by more than 70%, without affecting the inversion results. Numerical experiments with synthetic data show that reasonable fits can be obtained within a small number of iterations, with a few additional iterations required to eliminate unnecessary structure and find the model with minimum norm.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Magnetotellurics; Data-space method; 3D inversion; Occam's inversion

1. Introduction

The capability for routine three-dimensional inversion is a requirement for further progress with the magnetotelluric (MT) method, since two-dimensional (2D) interpretations frequently cannot explain important features present in field data sets from geologically

complex regions. There have been many recent efforts to develop 3D MT inversion algorithms, using a fairly broad range of approaches (e.g., Mackie and Madden, 1993; Zhdanov et al., 2000a; Newman and Alumbaugh, 2000; Farquharson et al., 2002). These schemes have been shown to recover conductivity variations reasonably well, at least for simple test cases with synthetic data. However, the 3D MT inverse problem is far from solved. Routine application of 3D inversion is still hindered by the requirement of a high-end workstation or

* Corresponding author. Tel.: +662 201 5770; fax: +662 354 7159.
E-mail address: scwsp@mahidol.ac.th (W. Siripunvaraporn).

a parallel machine, and issues of robustness and reliability with real data have barely been addressed for any of the proposed methods. Improvements in algorithms for routine 3D inversion are highly desirable.

One promising approach to improve efficiency is based on use of rapid approximate modeling schemes, such as the quasi-linear or quasi-analytic approximations (Torress-Verdin and Habashy, 1994; Zhdanov and Fang, 1996a; Tseng et al., 2003). Because the model response in these approximations is quasi-linear with respect to the modified conductivity tensor, simplified inversion methods are possible (e.g., Zhdanov and Fang, 1996b; Zhdanov et al., 2000b). Depending on the degree of simplification used, schemes like this can produce images of earth structure very rapidly. However, such approximate methods have their limitations, e.g., they work best when conductivity contrasts are low, and the general reliability and accuracy of these inversions are open to question. Although these rapid schemes are of unquestionable value, methods based on full solution of the EM induction equations will still be needed for many purposes.

Most 3D MT inversions have taken this more traditional inversion approach, based on minimizing a functional penalizing both misfit to the data and model “roughness” (e.g., Parker, 1994). To minimize the penalty functional gradients are computed, model parameters are adjusted, and the full forward problem is solved numerically to assess model fit. For example, Mackie (personal communication, 2002) has extended the non-linear conjugate gradients (NLCG; Rodi and Mackie, 2001) method to 3D, with the help of message passing interface (MPI) running on PC-clusters. Newman and Alumbaugh (2000) have applied a similar technique to 3D inversion, but on a massively parallel system. Sasaki (2001) and Farquharson et al. (2002) have developed 3D inversions based on a Gauss–Newton (GN) method. All of these schemes can be classified as model-space inversions, in the sense that search for the optimal conductivity is conducted in the M dimensional model parameter space.

With a model-space method, both the number of operations and required random access memory (RAM) depend strongly on the number of model parameters M . In particular, the most straightforward approach to penalty functional minimization in the model space (e.g., GN) requires forming and solving an $M \times M$ system of linear normal equations. For 3D inversion, M can

easily be so large that this would be impossible, both in terms of computational time and, especially, RAM. One way around this difficulty (e.g., Sasaki, 2001; Newman and Alumbaugh, 2000) is to use a very coarse parameterization of the conductivity variations, so that M remains small. But then results will depend strongly on the choice of model parameters, and unless one has strong a priori constraints on earth structure, results of such an inversion may be very misleading. Iterative inversion methods, such as conjugate gradient (CG; Mackie and Madden, 1993) or NLCG (Newman and Alumbaugh, 2000), avoid explicitly forming and storing the $M \times M$ coefficient matrix for the normal equations, and this can allow for more general and geologically realistic model parameterization. This general iterative approach has come to be seen as the only practical computational scheme for over-parameterized minimum structure 3D inversion where M is large.

In this paper, we present results from a newly developed data-space 3D MT inversion, in which minimization of the penalty functional is conducted in the N dimensional data space. With a data-space approach the null space (i.e., the part of the parameter space that has no effect on the data) is eliminated at the outset, and the $M \times M$ system of normal equations is replaced by an $N \times N$ system. Thus, the size of all computations and required arrays depends primarily on the number of independent data N , which for 3D geologically realistic modeling will generally be much less than M . Data-space methods have in fact been widely applied to inverse problems in geophysics (e.g., Parker, 1994) and other fields (Egbert et al., 1994; Chua and Bennett, 2001). The data-space approach allows us to consider inversion algorithms other than the CG approach, without severe restrictions on model parameterizations. The approach we explicitly consider here is a data-space variant on the Occam scheme. The original Occam’s inversion (Constable et al., 1987; deGroot-Hedlin and Constable, 1990) method was formulated in the model space. Here, we reformulate Occam in the data space, as in Siripunvaraporn and Egbert (2000) for the 2D MT inverse problem. As discussed in Siripunvaraporn and Egbert (2000), data-space algorithms can be made significantly more efficient by using a reduced basis approach. In this paper, we only consider implementation of the most straightforward data-space variant on Occam, as a test of the practicality of a data-space approach.

2. Occam's inversion

Occam's inversion seeks the “smoothest”, or minimum norm, model subject to an appropriate fit to the data (Constable et al., 1987). Mathematically, this objective is achieved by finding stationary points of an unconstrained functional $U(\mathbf{m}, \lambda)$:

$$U(\mathbf{m}, \lambda) = (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0) + \lambda^{-1} \{(\mathbf{d} - \mathbf{F}[\mathbf{m}])^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{F}[\mathbf{m}]) - X^{*2}\}. \quad (1)$$

Here $\mathbf{m}$ is the resistivity model, $\mathbf{m}_0$ the prior model, $\mathbf{C}_m$ the model covariance matrix which defines the model norm, $\mathbf{d}$ the observed data, $\mathbf{F}[\mathbf{m}]$ the model response, $\mathbf{C}_d$ the data covariance matrix, X^* the desired level of misfit, and λ^{-1} a Lagrange multiplier. Assuming the data covariance is correctly specified, X^* should in theory be one (Constable et al., 1987), and we use this value for our experiments with synthetic data. With real data X^* must generally be set to a somewhat higher level.

To find the stationary points of (1), instead of working with (1) directly, we differentiate the penalty functional $W_\lambda(\mathbf{m})$:

$$W_\lambda(\mathbf{m}) = (\mathbf{m} - \mathbf{m}_0)^T \mathbf{C}_m^{-1} (\mathbf{m} - \mathbf{m}_0) + \lambda^{-1} \{(\mathbf{d} - \mathbf{F}[\mathbf{m}])^T \mathbf{C}_d^{-1} (\mathbf{d} - \mathbf{F}[\mathbf{m}])\}, \quad (2)$$

with respect to $\mathbf{m}$. When λ is fixed, both U and W_λ have the same stationary points. By minimizing W_λ with a series of λ , the stationary points of U can be obtained (i.e., λ can be found such that the data misfit is X^{*2}).

2.1. Model-space method

Because of non-linearity of the MT inverse problem, an iterative approach is required, based on linearizing $\mathbf{F}[\mathbf{m}]$ such that:

$$\mathbf{F}[\mathbf{m}_{k+1}] = \mathbf{F}[\mathbf{m}_k + \Delta \mathbf{m}] = \mathbf{F}[\mathbf{m}_k] + \mathbf{J}_k (\mathbf{m}_{k+1} - \mathbf{m}_k). \quad (3)$$

Here the subscript k denotes iteration number, and $\mathbf{J}_k = (\partial \mathbf{F} / \partial \mathbf{m})_k$ is the $N \times M$ sensitivity matrix calculated at $\mathbf{m}_k$. Substituting (3) into (2), and finding the stationary points, we obtain a series of iterative ap-

proximate solutions:

$$\mathbf{m}_{k+1}(\lambda) = [\lambda \mathbf{C}_m^{-1} + \mathbf{\Gamma}_k^m]^{-1} \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{X}_k + \mathbf{m}_0, \quad (4)$$

where $\mathbf{X}_k = \mathbf{d} - \mathbf{F}[\mathbf{m}_k] + \mathbf{J}_k (\mathbf{m}_k - \mathbf{m}_0)$, and the “model-space cross-product” matrix $\mathbf{\Gamma}_k^m = \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{J}_k$ is an $M \times M$ positive semi-definite symmetric matrix.

To reach the ultimate goal of finding the stationary points of (1), in each iteration (4) is computed with a series of trial values of λ to minimize the misfit among solutions of this form. The goal of iterations in the early stage (Phase I) is to bring the misfit down to the target level, X^{*2} . Once the misfit reaches the desired level, the next stage begins by keeping the misfit at the desired level, varying λ to seek the model of smallest norm (Phase II) achieving the target misfit. For a variety of reasons, one may never reach the target misfit. An improvement of misfit from iteration to iteration can be expected, but even this is not guaranteed.

2.2. Data-space method

As shown by Parker (1994), and summarized in Appendix A, the solution for iteration k can be expressed as a linear combination of rows of the smoothed sensitivity matrix $\mathbf{C}_m \mathbf{J}_k^T$, i.e.

$$\mathbf{m}_{k+1} - \mathbf{m}_0 = \mathbf{C}_m \mathbf{J}_k^T \boldsymbol{\beta}_{k+1}, \quad (5)$$

where $\boldsymbol{\beta}_{k+1}$ is an unknown expansion coefficient vector of the basis functions $[\mathbf{C}_m \mathbf{J}_k^T]_j$; $j = 1, \dots, N$. Substituting (5) into a linearized form of (2), and solving for its stationary point, we again obtain a series of iterative solutions:

$$\boldsymbol{\beta}_{k+1} = [\lambda \mathbf{C}_d + \mathbf{\Gamma}_k^n]^{-1} \mathbf{X}_k, \quad (6)$$

where $\mathbf{\Gamma}_k^n = \mathbf{J}_k \mathbf{C}_m \mathbf{J}_k^T$ is the $N \times N$ “data-space cross-product” matrix, which is symmetric and positive semi-definite. Similarly to the standard model-space Occam's inversion, we can solve for $\boldsymbol{\beta}_{k+1}$ using (6), update the model, and then compute the misfit. All of these calculations are done with various values of λ for both Phases I and II, just as in the model-space approach.

The solutions obtained from both approaches, i.e., from (4) for the model-space method and from (6) and (5) for the data-space method, should in theory be identical if all parameters used are the same. The major difference between (4) and (6) is that the dimension of the system of equations to be solved can be significantly

reduced, from $M \times M$ in the model-space method, to $N \times N$ in the data-space method. In very many practical cases, N will be much less than M , especially for the 3D MT inversion problem we consider here. This reduction means a great saving on computational costs of both memory and CPU time.

Another distinction between the two methods is that the model covariance $\mathbf{C}_m$ is directly required for the data-space method, while its inverse is required in the model-space method. In the model-space method, the inverse of the model covariance ($\mathbf{C}_m^{-1}$) is, for reasons of efficiency, usually implemented as a sparse model roughness operator (Constable et al., 1987; deGroot-Hedlin and Constable, 1990). The exact inverse of a specific roughness operator ($\mathbf{C}_m$) cannot be determined in practice, because of the size of this matrix, which in general will be full. Conversely, a model covariance matrix $\mathbf{C}_m$, which is efficient for data-space computations (e.g., Siripunvaraporn and Egbert, 2000) cannot be inverted both due to size and ill-conditioning. It is thus impractical to compare directly model and data-space calculations with identical model covariances, for even a small model grid. Note that prior information, such as faults or an ocean, can be readily included in the model covariance $\mathbf{C}_m$, which is another advantage of the data-space method.

For further details on the Occam's inversion in both the model and data space, see Parker (1994), Siripunvaraporn and Egbert (2000), Constable et al. (1987), and deGroot-Hedlin and Constable (1990).

3. Inversion algorithm details

Our implementation of the 3D data-space Occam's inversion closely follows the 2D inversion described by Siripunvaraporn and Egbert (2000). Here, we briefly summarize algorithm details specific to the 3D implementation.

3.1. Three-dimensional forward modeling

Inversion relies heavily on forward modeling, both for computing model responses, and sensitivities. An efficient and accurate forward modeling code is thus essential (e.g., Mackie et al., 1994; Smith, 1996; Newman and Alumbaugh, 2000; Siripunvaraporn et al., 2002; Avdeev et al., 2002). We solve the second

order Maxwell's equations with a staggered grid finite difference numerical approximation. This approach to 3D forward modeling is flexible and allows large and complicated model structures (depending on computer resources) in an efficient way.

There are two forms of the second order Maxwell's equation: in terms of electric fields:

$$\nabla \times \nabla \times \mathbf{E} = i\omega\mu\sigma\mathbf{E}, \quad (7a)$$

or in terms of magnetic fields

$$\nabla \times \rho \nabla \times \mathbf{H} = i\omega\mu\mathbf{H}. \quad (7b)$$

Here μ is the air magnetic permeability, ω the angular frequency, σ the conductivity (the inverse of resistivity, ρ), $\mathbf{E}$ the electric field, and $\mathbf{H}$ the magnetic field. Siripunvaraporn et al. (2002) have shown that solutions obtained from equations formulated in terms of the electric fields (7a), with a staggered grid finite difference are less sensitive to grid resolution than those obtained from the magnetic formulation (7b). We therefore use (7a) rather (7b) for modeling.

With a staggered grid finite difference approximation to (7), we obtain the discrete system of equations $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{b}$ is the boundary electric fields, $\mathbf{x}$ the interior unknown electric fields, and $\mathbf{A}$ the symmetric coefficient matrix (but $\mathbf{A}$ is not Hermitian; it is complex only on the diagonal). The linear system of equations is then solved via the quasi-minimum residual (QMR) method, with a preconditioner formed by an incomplete LU decomposition of the diagonal sub-matrix of $\mathbf{A}$ (Siripunvaraporn et al., 2002). A divergence correction, similar to Smith (1996), is also applied to speed up convergence. The iterative solution is terminated once the level of normalized misfit, $\mathbf{r} = \|\mathbf{Ax} - \mathbf{b}\|/\|\mathbf{b}\|$, is below 10^{-8} . After solving for the interior electric fields, the magnetic fields at the surface are then computed and interpolated in the usual way (e.g., via the first order Maxwell's equations).

3.2. Data and model responses for 3D inversion

Two polarizations, $\mathbf{E}_x - \mathbf{H}_y$ and $\mathbf{E}_y - \mathbf{H}_x$, are computed via the forward modeling algorithm to generate the model responses. Each polarization has its own electric and magnetic fields. The fields from both

polarizations are related via the impedance tensor, $\mathbf{Z}$:

$$\begin{bmatrix} \mathbf{E}_x^1 & \mathbf{E}_x^2 \\ \mathbf{E}_y^1 & \mathbf{E}_y^2 \end{bmatrix} = \begin{bmatrix} \mathbf{Z}_{xx} & \mathbf{Z}_{xy} \\ \mathbf{Z}_{yx} & \mathbf{Z}_{yy} \end{bmatrix} \begin{bmatrix} \mathbf{H}_x^1 & \mathbf{H}_x^2 \\ \mathbf{H}_y^1 & \mathbf{H}_y^2 \end{bmatrix}. \quad (8)$$

Here $\mathbf{E}_x^1$ and $\mathbf{E}_x^2$ are the x -component of electric fields for the $\mathbf{E}_x - \mathbf{H}_y$ and $\mathbf{E}_y - \mathbf{H}_x$ polarizations, respectively, and similarly for other field components. Note that to calculate any parts of the impedance tensor, fields from both polarizations are required.

Model responses for 2D MT inversions are usually the apparent resistivities and phases calculated from the off-diagonal terms of $\mathbf{Z}$. However, for the 3D case, the diagonal terms ($\mathbf{Z}_{xx}$ and $\mathbf{Z}_{yy}$) can also become significant, and should be included in the inversion. In our algorithm, we invert the full impedance tensor, $\mathbf{Z}$, including both real and imaginary parts.

3.3. Sensitivity matrix

The sensitivity calculation is essential to our inversion approach. We follow the general approach described in Siripunvaraporn and Egbert (2000) or Rodi (1976), using reciprocity. For the 3D case, as we invert the full impedance tensor, the sensitivity calculation at each station (and for each period) requires solving two forward problems, one for each polarization. For example, computing $\partial \mathbf{Z}_{xy} / \partial \mathbf{m}$ requires calculating $\partial \mathbf{E}^1 / \partial \mathbf{m}$ and $\partial \mathbf{E}^2 / \partial \mathbf{m}$, and likewise for other sensitivity terms. This is not the case for 2D where only one forward calculation is required for each sensitivity term. Our mathematical derivation for the sensitivity matrix of the MT impedance tensor is very similar to that presented in Newman and Alumbaugh (2000).

3.4. Model covariance

The model covariance matrix $\mathbf{C}_m$ characterizes the expected magnitude and smoothness of resistive variation relative to the base model. Here, a model covariance similar to that used by Siripunvaraporn and Egbert (2000) for the 2D case is applied and extended to the 3D case. For the data-space method, the model covariance itself is never constructed. Only the product with the sensitivity matrix, $\mathbf{C}_m \mathbf{J}^T$, is required. For a model covariance with a Gaussian correlation function the product of $\mathbf{C}_m$ with any model vector $\mathbf{a}$ can be computed by

solving a diffusion equation with initial conditions $\mathbf{a}$ (Egbert et al., 1994; Siripunvaraporn and Egbert, 2000).

To avoid solving the 3D diffusion equation, we instead solve 1D diffusion equations alternatively between vertical and horizontal directions (both x - and y -directions). Our approach can be viewed as a simple operator splitting solution (e.g., Press et al., 1992) of the 3D diffusion equation. The decorrelation scale for the diffusion equation in all directions varied in space, with length scales proportional to $\sqrt{(4\delta\tau)}$ of the local grid resolution, where δ (between 0 and 1) and τ are given by users. For the two example cases presented below, values of δ are 0.2 and 0.1, respectively, while τ is 10 for both cases. Alternative approaches for selecting decorrelation length scales are discussed in Siripunvaraporn and Egbert (2000).

4. Synthetic data example and discussions

To test the 3D data-space algorithm, we have run the inversion program on two synthetic data sets. All computations are performed on a Dec Alpha 666 MHz machine with 1 Gbyte of RAM, so these computations could easily be reproduced on a common modern PC. However, the size of the data and model used in these inversion tests are rather limited. Faster computers with more RAM will be required for more realistic applications.

4.1. Synthetic case I

The first data set is generated from a very simple model (Fig. 1), consisting of a conductive block of $1 \Omega\text{m}$ ($16 \text{ km} \times 16 \text{ km} \times 5 \text{ km}$) buried 100 m beneath the surface of a $100 \Omega\text{m}$ half space. Data for 36 sites, distributed as shown in Fig. 1 as solid dots, were generated by solving (7a) on $56 \times 56 \times 28$ (+7 air layers) grid. The complex impedance tensor ($\mathbf{Z}_{xx}$, $\mathbf{Z}_{xy}$, $\mathbf{Z}_{yx}$ and $\mathbf{Z}_{yy}$) for five periods (0.1, 1, 10, 100 and 1000 s) are inverted with 5% Gaussian noise. The data variance is assumed to be 5% of $|\mathbf{Z}_{xy}\mathbf{Z}_{yx}|^{1/2}$. The model mesh used for the inversion is $28 \times 28 \times 21$ (+7 air layers). Note that this discretization is different from that used to generate the data. In this simple test, the total number of data $N = 1440$, and the total number of model parameters $M = 16,464$.

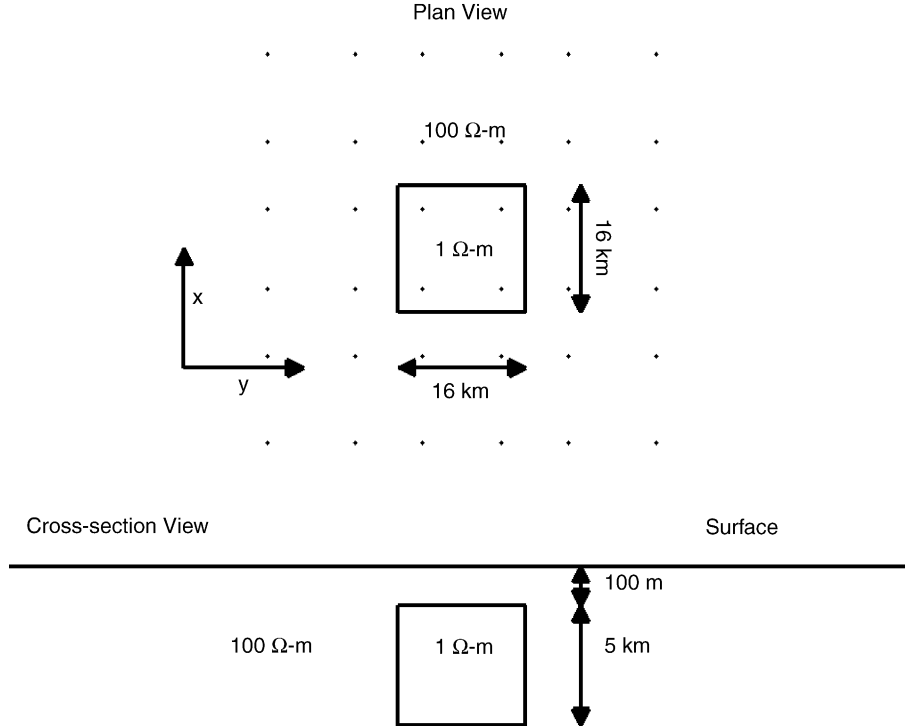


Fig. 1. A simple synthetic model used in the first test of the inversion. The solid dots indicate the observation sites. The cross-section in the lower panel picture is not to scale.

The inversion is started from a $50 \Omega\text{m}$ half space, which is also used as the base model, $\mathbf{m}_0$ for the inversion. The initial RMS for this starting model is about 23. Fig. 2 displays the convergence of the algorithm. For each iteration a series of λ values are used to search for the minimum RMS; these are marked by solid symbols. Note that the value of λ at the minimum changes significantly with iterations. The dotted and dashed lines give results for two variants, which we discuss below. Within three iterations, the inversion has converged to the desired level of misfit, i.e., completing Phase I. The inversion spends another two iterations to search for the model with minimum norm (Phase II), while keeping the misfit at the target misfit. The program is terminated once the model norm changes only slightly. For Phase II, λ is slightly perturbed to a higher value, and unnecessary structures are removed from the model.

For each iteration, the required calculation time is about 17 h (for the dashed lines of Fig. 2), for a total of 84 h for five iterations. More than 90% of the computing time is used in constructing the sensitivity matrix,

with most of this time spent solving forward problems. To construct the sensitivity matrix, the number of forward solutions per iteration is about $2N_s \times N_p$, where N_p is the number of periods, and N_s is the number of station. Reducing the time required for solving the forward problem when constructing the sensitivity matrix would significantly speed up the inversion. To achieve accuracy in forward computations, we usually terminate the QMR iterations when the normalized residual is less than 10^{-8} . For our first tests with the inversion (dashed lines in Fig. 2), we used the same convergence criterion for computation of the sensitivities. We will refer to this sensitivity calculation as “the full convergence case”.

In an effort to reduce computational time, we set the termination condition to a minimum number of 40 iterations, and a normalized residual of 10^{-4} . This requires about 1/3 to 1/4 of the iterations used for the full convergence case. We will refer to this as “the relaxed convergence case”. The solid lines of Fig. 2 display the convergence of the relaxed convergence case, and

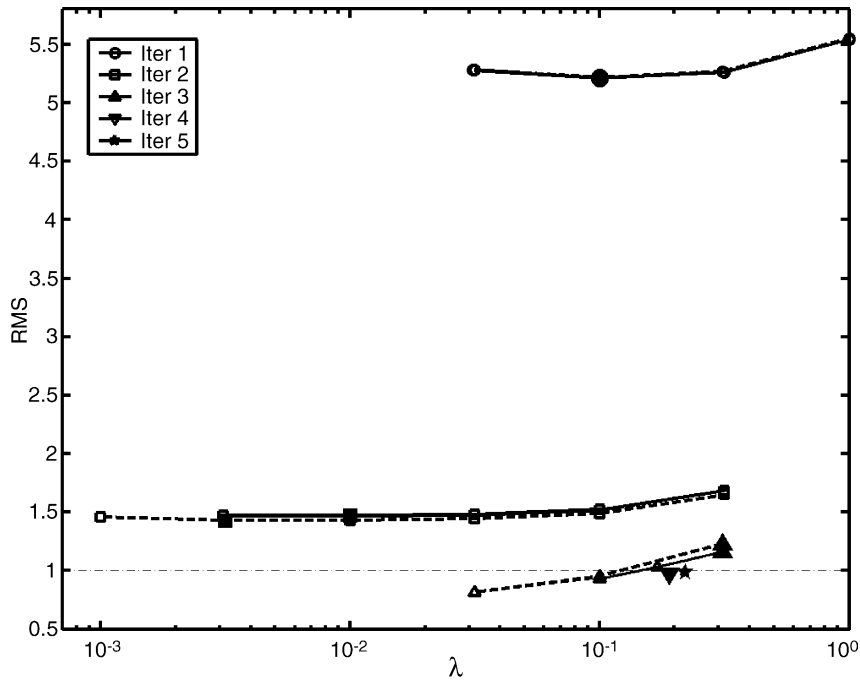


Fig. 2. Plots of RMS vs. λ for each iteration for inverting synthetic data generated from the model of Fig. 1. The dashed lines indicate the full convergence scheme, while the solid lines indicate a relaxed convergence scheme, when computing for the sensitivities. For most outer-loop iterations, the two lines appear to be on top of each other. Solid marks indicate the minimum RMS value for each iteration. Iterations 1–3 indicate Phase I, and iterations 4 and 5 (plotted only at the desired misfit) display Phase II of the inversion.

shows that using this scheme does not have much effect on the outcome of the inversion. However, the computational time per iteration is significantly reduced to about 5.8 h per iteration (about 29 h for five iterations) or only about 30% of the total calculation time with the full convergence case. Tests on other synthetic data sets yielded similar results.

Fig. 3 displays an inverse solution obtained at the target misfit with the minimum model norm using the relaxed convergence case after the fifth iteration. The upper panels of Fig. 3 displays the plan view at the surface, at 200 m depth and at 3 km depth, while the lower panel displays the cross-section view through the middle of the conductive block at $X=0$ km. The plan view at the surface reflects the thin resistive layer on top of the conductive block, while at 200 m and 3 km depth the conductive block is clearly seen. Fig. 3 shows that the inversion recovers the conductivity and position of the block. The image is not perfect near the block edges, primarily due to the limited number of sites. Additional short period data (e.g., lower than

0.1 s) would undoubtedly improve resolution of the thin resistive surface layer. In addition, structures near the bottom of the conductive block are not well resolved. Again, longer period data may help further constrain the structure at depth. However, it is well-known that the bottom of a conductive layer is difficult to recover with MT data. This difficulty is intrinsic to induction data, and is not specific to our inversion approach; e.g., see Siripunvaraporn and Egbert (2000) or Newman and Alumbaugh (2000), among many others for examples. As suggested by Fig. 2, the inverted model from the full convergence case is very similar to the model of Fig. 3.

In comparison with the model-space method, computing requirements for building the sensitivity matrix and for computing the responses and misfit are about the same. The main difference lies in solving the system of Eqs. (6) or (4) for each λ value. For this synthetic case, the coefficient matrices are of size 1440×1440 and $16,464 \times 16,464$, respectively. Solving the $16,464 \times 16,464$ would require a factor of 11^3

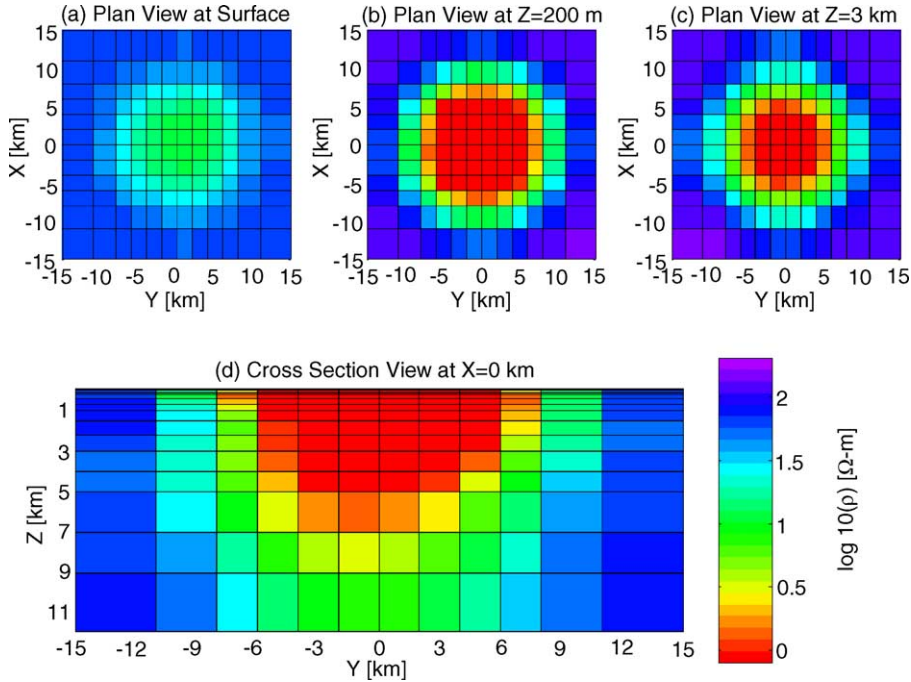


Fig. 3. An inverted model after the fifth iterations with an RMS value at 1, from synthetic data in Fig. 1. The top panels (a)–(c) is a plan view at the surface, at 200 m and at 3 km depth, and the bottom panel (d) is a cross section view cutting across the conductive block at $X = 0$ km. The solution is shown only in the central area around the anomalies, not in the full 28×28 model domain.

more operations. Since the system must be solved for each λ (an average of four per iteration for Phase I, and 3 for Phase II), this would push the computational time per iteration to an unacceptable level. In addition, storage of the coefficient matrix would require a factor of 130 times more memory than is required by the data-space method, making a model-space Occam's inversion impractical unless the model is severely under parameterized (Sasaki, 2001). Other model-space approaches, which do not require computation or storage of full sensitivity matrix, such as the NLCG scheme of Newman and Alumbaugh (2000), would perhaps be more practical. However, to find a minimum structure model subject to specified data misfit with such approaches multiple inversion runs with different values of λ would be required.

4.2. Synthetic case II

We also tested the inversion on a more complex model, one that is similar to a model previously used in many 3D forward modeling studies (see Wannamaker,

1991; Mackie et al., 1994; Siripunvaraporn et al., 2002; Avdeev et al., 1997). The model consists of resistive and conductive blocks buried in a two-layered Earth (Fig. 4). The observation sites are shown as solid dots in Fig. 4 for 40 sites, in a 5×8 grid. Data for 5 periods at 0.1, 1, 10, 100 and 1000 s are inverted. The full complex impedance tensor is inverted, so the total number of data is $N = 1600$. The data are added with 5% noise. The mesh used in the inversion is discretized at 21 blocks in x , 28 blocks in y , and 21 blocks in z (with seven air layers on top), or $M = 12,348$. Note that a different mesh (at $42 \times 56 \times 42$) is used to generate the synthetic data. The initial model and also the base model ($\mathbf{m}_0$) for the inversion is the same, i.e., a 50 Ωm half space.

The convergence plot is shown in Fig. 5, and the inverted model is shown in Fig. 6 (after the sixth iteration). The inversion spends roughly 6.5 h per iteration, using the relaxed convergence scheme when computing the sensitivity matrix. From Fig. 5, the inversion has reached the desired RMS within three iterations (Phase I), and spends another three iterations to search for the

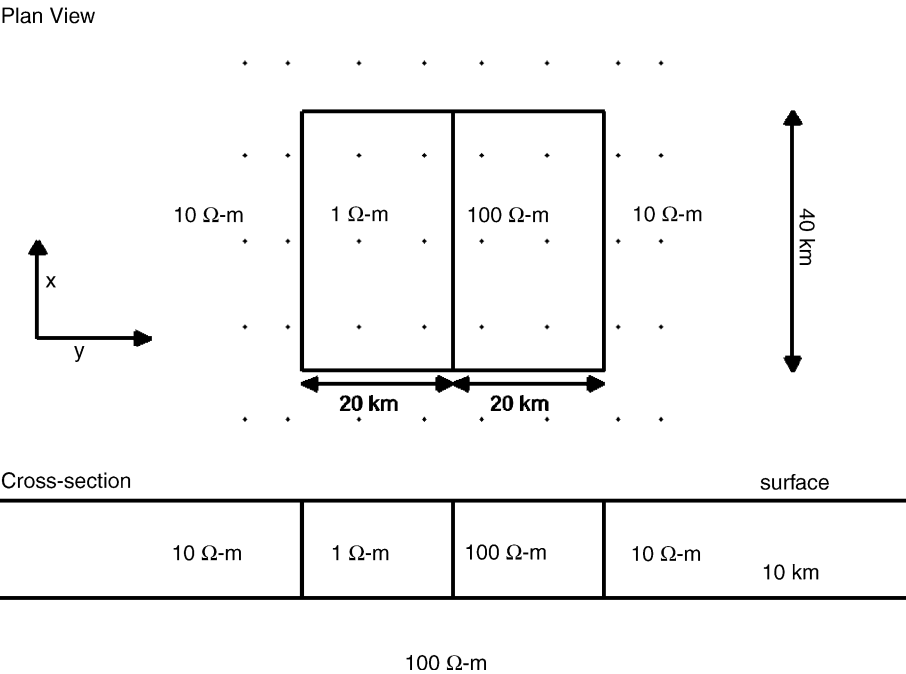


Fig. 4. Another synthetic model used to test the inversion. A cross-section view is a profile cutting across the middle of the conductive block and resistive block in the plan view (upper panel). The solid dots indicate the observation sites. A lower panel picture is not on-scale.

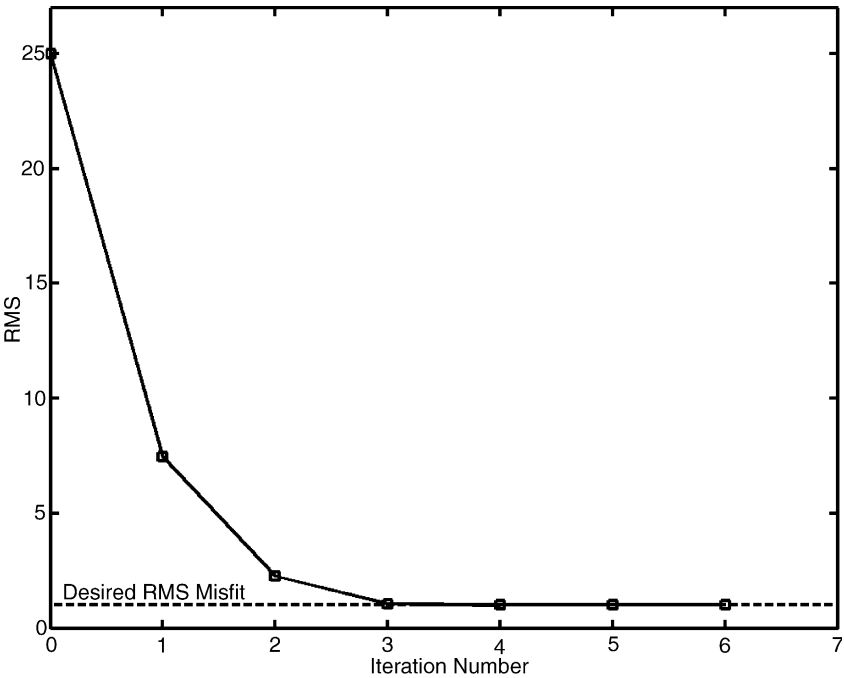


Fig. 5. RMS plots vs. iteration number for the inversion of the synthetic data generated from model of Fig. 4.

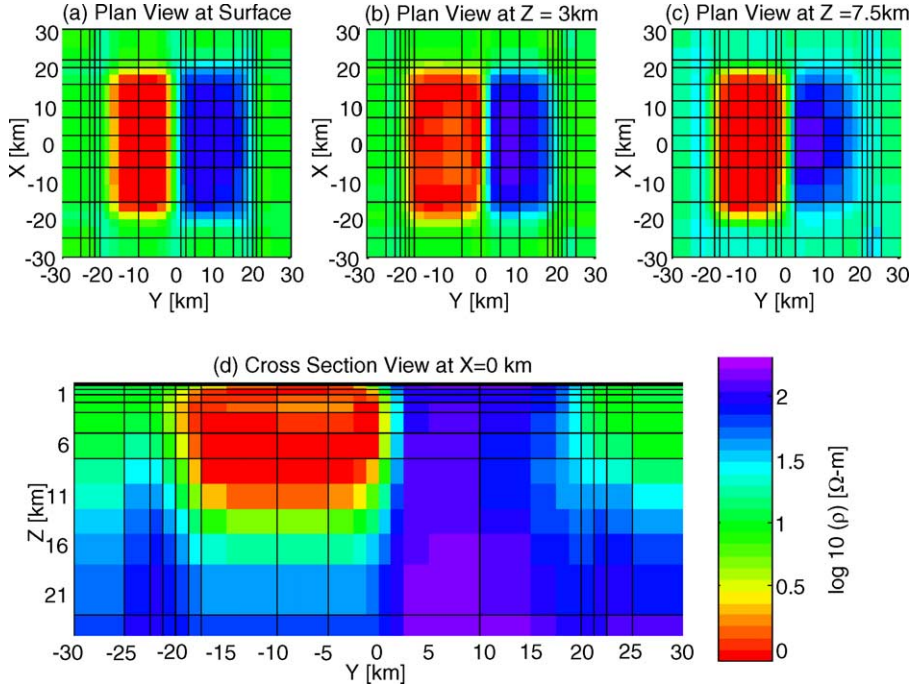


Fig. 6. An inverted model after the sixth iterations, with an RMS value at 1 from synthetic data in Fig. 4. The top panels (a)–(c) is a plan view at the surface, at 3 km and at 7.5 km depth, and the bottom panel (d) is a cross section view cutting across the conductive block at $X = 0$ km. As in Fig. 3 the solution is shown only in the central part of the model domain.

model with minimum structure (Phase II). The inversion can recover both conductive and resistive bodies from the $10 \Omega\text{m}$ background host (Fig. 6), even though not perfect near the edges and at great depths. Similar to case I, if this inversion were performed in the model space, it would require a huge amount of both computational time and RAM, which would make it impractical to run on a regular PC.

5. Discussions and conclusions

A three-dimensional MT inversion algorithm based on the Occam approach has been developed to seek the smoothest model subject to an appropriate fit to the data. By transforming from model space to data space its computational costs are significantly reduced, allowing us to invert modest 3D MT data sets on a personal computer (PC) in a relatively short time.

In the two examples presented here $N \ll M$, the efficiency advantage of the data-space method is evident. More generally it is possible that N could be equal or

larger than M , for example if we inverted many more frequencies in our test problems. In this case, the data-space method would not be superior to the model space in term of the computational costs. However, as shown in Siripunvaraporn and Egbert (2000), data redundancy (which is manifested in the near linear dependence of the columns of $\mathbf{C}_m \mathbf{J}^T$) implies that the effective value of N can always be significantly reduced by restricting trial solutions to a subspace of the columns of $\mathbf{C}_m \mathbf{J}^T$. Such data sub-space methods guarantee significant computational savings regardless of the values of N and M .

The data-space approach we have discussed here is not limited to an Occam (or reduced basis Occam) approach. Other inversion approaches may also be combined with or adapted to a data-space formulation. For example, it may not be necessary to compute and store the full sensitivity and cross-product matrices. A conjugate gradient method, in which multiplication of the sensitivity matrix by an arbitrary vector is achieved by solving several forward problems (without computing or storing the matrix of sensitivities; Mackie and

Madden, 1993) could also be implemented in the data space. Such a data-space conjugate gradients scheme would allow solution of much larger inverse problems. Compared to a model-space CG inversion, the system of equations to be solved would be significantly reduced by transformation to the data space. In oceanographic data assimilation problems a data-space CG approach has proven to be more stable and effective than comparable model-space approaches (Bennett et al., 1997; Egbert, 1997; Chua and Bennett, 2001), so a data-space variant on the CG approach to MT inversion deserves consideration.

Note that the CG and NLCG schemes only really make sense in the case where λ is fixed. For an Occam algorithm, where the normal equations of (4) or (6) are solved a number of times in each linearization step (to search for the minimum misfit or minimum norm model) a CG approach probably would not be practical or efficient. Thus, the data-space formulation is probably the only practical way to implement the Occam scheme in 3D. Furthermore, hybrid approaches which combine elements of the data-space method, CG, Occam, and other approaches are possible, and should be explored. The data-space formulation discussed here should be kept in mind as we explore efficient and robust solutions to the 3D MT inverse problem.

Acknowledgments

This research has been supported by a grant from the Thailand Research Fund (TRF:PDF/37/2543) to W.S., and by DOE-FG0302ER15318 to G.E., and partly by a JSPS postdoctoral fellowships (P02053) and a grant from MEXT, Japan (02053) to W.S. and M.U. W.S. thanks the Faculty of Graduate Studies and Faculty of Science, Mahidol University for travel support to attend the XVI Induction Workshop in Santa Fe, USA, and also the Induction Workshop Committee for waiving the registration fee. The authors would like to thank Dimitry Avdeev and Yuji Mitsuhata for their comments to help improving the manuscript.

Appendix A

From Eq. (4), it follows that

$$\begin{aligned}
 \mathbf{m}_{k+1}(\lambda) - \mathbf{m}_0 &= [\lambda \mathbf{C}_m^{-1} + \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{J}_k]^{-1} \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{X}_k \\
 &= \mathbf{C}_m [\lambda \mathbf{I} + \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{J}_k \mathbf{C}_m]^{-1} \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{X}_k \\
 &= \mathbf{C}_m [\lambda \mathbf{I} + \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{J}_k \mathbf{C}_m]^{-1} (\mathbf{J}_k^{-T})^{-1} \mathbf{C}_d^{-1} \mathbf{X}_k \\
 &= \mathbf{C}_m [\mathbf{C}_d \mathbf{J}_k^{-T} (\lambda \mathbf{I} + \mathbf{J}_k^T \mathbf{C}_d^{-1} \mathbf{J}_k \mathbf{C}_m)]^{-1} \mathbf{X}_k \\
 &= \mathbf{C}_m \mathbf{J}_k^T [\lambda \mathbf{C}_d + \mathbf{J}_k \mathbf{C}_m \mathbf{J}_k^T]^{-1} \mathbf{X}_k \\
 &= \mathbf{C}_m \mathbf{J}_k^T \boldsymbol{\beta}_{k+1}
 \end{aligned} \tag{A1}$$

where $\boldsymbol{\beta}_{k+1} = [\lambda \mathbf{C}_d + \mathbf{J}_k \mathbf{C}_m \mathbf{J}_k^T]^{-1} \mathbf{X}_k$. We therefore obtain Eq. (5). In derivation presented in (A1), we used properties: $(\mathbf{J}_k^{-T})^{-1} = \mathbf{J}_k^T$ and $(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$.

References

- Avdeev, D.B., Kuvshinov, A.V., Pankratov, O.V., Newman, G.A., 2002. Three-dimensional induction logging problems. Part 1. An integral equation solution and model comparisons. *Geophysics* 67, 413–426.
- Avdeev, D.B., Kuvshinov, A.V., Pankratov, O.V., Newman, G.A., 1997. High-performance three-dimensional electromagnetic modelling using modified Neumann series. Wide-band numerical solution and examples. *J. Geomag. Geoelectr.* 49, 1519–1539.
- Bennett, A.F., Chua, B.S., Leslie, L.M., 1997. Generalized inversion of a global numerical weather prediction model. *Meteorol. Atmospheric Phys.* 60, 165–178.
- Chua, B.S., Bennett, A.F., 2001. An inverse ocean modeling system. *Ocean Model.* 3, 137–165.
- Constable, C.S., Parker, R.L., Constable, C.G., 1987. Occam's inversion: a practical algorithm for generating smooth models from electromagnetic sounding data. *Geophysics* 52, 289–300.
- deGroot-Hedlin, C., Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics* 55, 1613–1624.
- Egbert, G.D., Bennett, A.F., Foreman, M.G., 1994. TOPEX/POSEIDON tides estimated using a global inverse model. *J. Geophys. Res.* 99, 24821–24852.
- Egbert, G.D., 1997. Tidal data inversion: interpolation and inference. *Progr. Oceanogr.* 40, 53–80.
- Farquharson, C.G., Oldenburg, D.W., Haber, E., Shekhtman, R., 2002. Three-dimensional forward-modeling and inversion algorithms for magnetotelluric data. In: *Proceedings of the 16th Workshop on Electromagnetic Data*, Santa Fe, New Mexico, June 16–22.
- Mackie, R.L., Madden, T.R., 1993. Three-dimensional magnetotelluric inversion using conjugate gradients. *Geophys. J. Int.* 115, 215–229.
- Mackie, R.L., Smith, J.T., Madden, T.R., 1994. Three-dimensional electromagnetic modeling using finite difference equations: the magnetotelluric example. *Radio Sci.* 29, 923–935.

- Newman, G.A., Alumbaugh, D.L., 2000. Three-dimensional magnetotelluric inversion using non-linear conjugate gradients. *Geophys. J. Int.* 140, 410–424.
- Parker, R.L., 1994. *Geophysical Inverse Theory*. Princeton University Press.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 1992. *Numerical Recipes in Fortran: the art of scientific computing*, 2nd ed. Cambridge University Press.
- Rodi, W.L., 1976. A technique for improving the accuracy of finite element solutions for magnetotelluric data. *Geophys. J. Roy. Astr. Soc.* 44, 483–506.
- Rodi, W.L., Mackie, R.L., 2001. Nonlinear conjugate gradients algorithm for 2D magnetotelluric inversion. *Geophysics* 66, 174–187.
- Sasaki, Y., 2001. Full 3D inversion of electromagnetic data on PC. *J. Appl. Geophys.* 46, 45–54.
- Siripunvaraporn, W., Egbert, G., 2000. An Efficient data-subspace inversion method for 2D magnetotelluric data. *Geophysics* 65, 791–803.
- Siripunvaraporn, W., Egbert, G., Lenbury, Y., 2002. Numerical accuracy of magnetotelluric modeling: a comparison of finite difference approximations. *Earth Planets Space* 54, 721–725.
- Smith, J.T., 1996. Conservative modeling of 3D electromagnetic fields. Part II. Biconjugate gradient solution and an accelerator. *Geophysics* 61, 1319–1324.
- Torress-Verdin, C., Habashy, T., 1994. Rapid 2.5-dimensional modelling and inversion via a new nonlinear scattering approximations. *Radio Sci.* 29, 1051–1071.
- Tseng, H.-W., Lee, K.H., Becker, A., 2003. 3D interpretation of electromagnetic data using a modified Born approximation. *Geophysics* 68, 127–137.
- Wannamaker, P.E., 1991. Advances in three-dimensional magnetotelluric modeling using integral equations. *Geophysics* 56, 1716–1728.
- Zhdanov, M.S., Fang, S., 1996a. Quasi-linear approximation in 3D electromagnetic modeling. *Geophysics* 61, 646–665.
- Zhdanov, M.S., Fang, S., 1996b. 3D quasi-linear electromagnetic inversion. *Radio Sci.* 31, 741–754.
- Zhdanov, M.S., Dmitriev, V.I., Fang, S., Hursan, G., 2000a. Quasi-analytical approximation and series in 3D electromagnetic modeling. *Geophysics* 65, 1746–1757.
- Zhdanov, M.S., Fang, S., Hursan, G., 2000b. Electromagnetic inversion using quasi-linear approximation. *Geophysics* 65, 1501–1513.