Minimum support nonlinear parametrization in the solution of a 3D magnetotelluric inverse problem

Michael Zhdanov and Ekaterina Tolstaya

Department of Geology and Geophysics, University of Utah, Salt Lake City, UT 84112, USA

Abstract

In this paper we describe a new approach to sharp boundary geophysical inversion. We demonstrate that regularized inversion with a minimum support stabilizer can be implemented by using a specially designed nonlinear parametrization of the model parameters. This parametrization plays the same role as transformation into the space of the weighted model parameters, introduced in the original papers on focusing inversion. It allows us to transform the nonquadratic minimum support stabilizer into the traditional quadratic minimum norm stabilizer, which simplifies the solution of the inverse problem. This transformation automatically ensures that the solution belongs to the class of models with a minimum support. The method is illustrated with synthetic examples of 3D magnetotelluric inversion for an earth conductivity structure. To simplify the calculations, in the initial stage of the iterative inversion we use the quasi-analytical approximation developed by Zhdanov and Hursán (2000 *Inverse Problems* **16** 1297–322). However, to increase the accuracy of inversion, we apply rigorous forward modelling based on the integral equation method at the final stage of the inversion. To obtain a stable solution of a 3D inverse problem, we use the Tikhonov regularization method with a new nonlinear parametrization. This technique leads to the generation of a sharp image of anomalous conductivity distribution. The inversion is based on the regularized conjugate gradient method.

1. Introduction

One of the very important problems in the inversion of geophysical data is developing a stable inverse problem solution, which, at the same time, can produce a sharp image of the target. Here we consider an ill-posed problem of reconstructing the inhomogeneous conductivity distribution of rock formations from the measured scattered electromagnetic (EM) field data. The traditional inversion methods are usually based on the Tikhonov regularization theory, which provides a stable solution of the inverse problem. Usually a maximum smoothness stabilizing functional is used to stabilize the inversion process. The obtained solution is a

smooth image, which in many practical situations does not describe the examined object properly.

The problem of reconstructing discontinuous images was most intensively researched in papers dedicated to image processing, image reconstruction and medical tomography (see, for example, Geman and Reynolds 1992, Geman and Yang 1995, Vogel 1997, Lobel *et al* 1997). For geophysical applications, this problem was investigated in Last and Kubik (1983), Portniaguine and Zhdanov (1999, 2002), Mehanee and Zhdanov (2002) and Zhdanov (2002). It was demonstrated in the cited publications that the images with sharp boundaries can be recovered by regularized inversion algorithms based on a new family of stabilizing functionals. Particularly, minimum support (MS) and minimum gradient support (MGS) functionals were found extremely effective in the solution of a geophysical inverse problem for mineral exploration (Zhdanov and Hursán 2000, Zhdanov 2002). These new stabilizers select inverse models within the class of models with the minimum volume of domain with anomalous parameter distribution (MS stabilizer) or with the minimum volume of area where the gradient of the model parameters is nonzero (MGS stabilizer). These classes of models describe compact objects (minimum support) or objects with sharp boundaries (minimum gradient support), which are typical targets in geophysical mineral exploration, for example.

Unfortunately, very often, these focusing stabilizing functionals are not convex, which complicates the minimization of the Tikhonov parametric functional. The nonconvexity of the stabilizing functional means that the functional may have several local extrema, where the first variation is equal to zero. The local minima slow down the convergence of the iterative schemes, which are usually used in minimization, and make it more difficult to find a global minimum.

It was demonstrated in Zhdanov (2002) that the nonquadratic (nonconvex) stabilizing functional can be represented in the form of a quadratic (convex) functional by a linear transformation of the original model parameters into the space of the weighted model parameters. This linear transformation is updated from iteration to iteration, which is equivalent to the re-weighting of the model parameters. In other words, we apply a set of linear transformations with the repeatedly updated weighting matrices to transform the nonquadratic functional into the Tikhonov quadratic stabilizer. For example, we can solve the inverse problem using the re-weighted conjugate gradient (CG) method with repeated modification of the model parameter weights after every few iterations.

In this paper we suggest using a nonlinear transformation instead of a set of variable linear transformations. In particular, we transform the nonquadratic minimum support stabilizing functional into a quadratic one by using a specially selected nonlinear transformation of the model parameters. We call this transformation *a minimum support nonlinear parametrization*. There are several advantages of this approach over the earlier approach of re-weighted minimum support minimization. First of all, with nonlinear parametrization, the parametric functional continuously decreases with the iteration number, which makes it easier to select the termination criterion for the iterative process, while in the framework of the previous approach this functional may increase after re-weighting. Second, there is no need to select the optimal number of re-weighting steps, as was the case with the original method. Finally, the convex nature of stabilizing functionals with the new parametrization ensures fast convergence of gradient-type iterative algorithms to the solution of the original inverse problem.

In geophysical inversion, we usually know *a priori* some physically meaningful bounds for the model parameters (the conductivity must be positive, for example). In order to introduce physical limits on the model parameters using new variables, it is necessary to impose upper and lower bounds on these variables during the inversion process. For this purpose we employ a gradient projection technique with a nonlinear projection operation.

We illustrate this new technique with synthetic examples of 3D magnetotelluric (MT) inversion for an earth conductivity structure. The numerical testing results indicate the effectiveness of this approach. The nonlinear parametrization can be used, however, for the solution of different geophysical inverse problems.

2. Principles of re-weighted regularized inversion

For completeness, we begin our paper with a short summary of the basic principles of re-weighted regularized inversion with the minimum support stabilizer. Let us consider a general discrete geophysical inverse problem, described by the operator equation

$$\mathbf{d} = \mathbf{A}(\mathbf{m}), \tag{1}$$

where, in general, $\mathbf{A}$ is a nonlinear vector operator, $\mathbf{m}$ represents the model parameters and $\mathbf{d}$ are observed geophysical data. We assume that N_d measurements are performed in some geophysical experiment. Then we can treat these values as the components of the N_d-dimensional vector $\mathbf{d}$. Similarly, some model parameters can be represented as the components of a vector $\mathbf{m}$ of order N_m:

$$\mathbf{d} = \left[d_1, d_2, d_3, \dots, d_{N_d} \right]^T,$$

$$\mathbf{m} = \left[m_1, m_2, m_3, \dots, m_{N_m} \right]^T,$$

where the superscript T denotes the transpose of the two vectors. Note that we assume that the model parameters are described by real numbers, while the components of vector $\mathbf{d}$ can be complex numbers.

The inverse problem (1) is usually ill-posed, i.e., the solution can be nonunique and unstable. We solve this ill-posed inverse problem using the regularization theory (Tikhonov and Arsenin 1977), which is based on the minimization of the Tikhonov parametric functional:

$$P^{\alpha}(\mathbf{m}) = \varphi(\mathbf{m}) + \alpha s(\mathbf{m}) \rightarrow \min, \tag{2}$$

where $\varphi(\mathbf{m})$ is a misfit functional between the theoretical values $\mathbf{A}(\mathbf{m})$ and the observed data $\mathbf{d}$, $s(\mathbf{m})$ is a stabilizing functional and α is a regularization parameter.

The misfit functional $\varphi(\mathbf{m})$ is usually selected in the complex Euclidean metric of data space as a weighted norm square of the difference between the observed and predicted data (errors):

$$\varphi(\mathbf{m}) = \|\widehat{\mathbf{W}}_d(\mathbf{A}(\mathbf{m}) - \mathbf{d})\|^2 = (\mathbf{A}(\mathbf{m}) - \mathbf{d})^* \widehat{\mathbf{W}}_d^2 (\mathbf{A}(\mathbf{m}) - \mathbf{d}), \tag{3}$$

where the asterisk '*' denotes a transposed complex conjugate matrix, $\mathbf{W}_d$ is the data weighting matrix which allows us to set the variance for each datum to its appropriate level.

The optimal value of α is determined from the misfit condition,

$$\varphi(\mathbf{m}) = \delta_d, \tag{4}$$

where δ_d is the noise level of the data.

The stabilizing functional $s(\mathbf{m})$ can be selected, for example, as a norm square of the difference between the current and *a priori* models:

$$s_{MN}(\mathbf{m}) = \|\mathbf{m} - \mathbf{m}_{\text{apr}}\|^2 = (\mathbf{m} - \mathbf{m}_{\text{apr}})^T (\mathbf{m} - \mathbf{m}_{\text{apr}}), \tag{5}$$

where the superscript 'T' denotes a transposed matrix. This is a minimum norm stabilizer which provides, usually, a relatively smooth image of the inverse model. Substituting (3) and (5) into (2), we arrive at the conventional Tikhonov parametric functional

$$P^{\alpha}(\mathbf{m}) = (\mathbf{A}(\mathbf{m}) - \mathbf{d})^* \mathbf{W}_d^2 (\mathbf{A}(\mathbf{m}) - \mathbf{d}) + \alpha(\mathbf{m} - \mathbf{m}_{\text{apr}})^T (\mathbf{m} - \mathbf{m}_{\text{apr}}) \rightarrow \min. \tag{6}$$

In order to generate a compact image of the geophysical model with sharp boundaries, Zhdanov and Hursán (2000), and Mehanee and Zhdanov (2002) applied a minimum support stabilizer, which is a nonquadratic functional of the form

$$s_{MS}(\mathbf{m}) = (\mathbf{m} - \mathbf{m}_{\mathrm{apr}})^T [(\widehat{\mathbf{m}} - \widehat{\mathbf{m}}_{\mathrm{apr}})^2 + e^2 \widehat{\mathbf{I}}]^{-1} (\mathbf{m} - \mathbf{m}_{\mathrm{apr}}), \tag{7}$$

where $\widehat{\mathbf{m}}$ and $\widehat{\mathbf{m}}_{\mathrm{apr}}$ are $N_m \times N_m$ diagonal matrices of inverse model parameters (current and *a priori*, respectively):

$$\widehat{\mathbf{m}} = \mathrm{diag}(m_1, m_2, \ldots, m_{N_m}), \qquad \widehat{\mathbf{m}}_{\mathrm{apr}} = \mathrm{diag}(m_{1\mathrm{apr}}, m_{2\mathrm{apr}}, \ldots, m_{N_m\mathrm{apr}}),$$

e is the focusing parameter and $\widehat{\mathbf{I}}$ is the $N_m \times N_m$ identity matrix. It was shown by Portniaguine and Zhdanov (1999) that this functional minimizes an area of nonzero parameter distribution (minimizes the support of the inverse model), if e tends to zero: $e \to 0$.

Following Zhdanov (2002) we note that this stabilizing functional can be expressed as a pseudo-quadratic functional of the model parameters,

$$s_{MS}(\mathbf{m}) = [\widehat{\mathbf{W}}_e(\mathbf{m} - \mathbf{m}_{\mathrm{apr}})]^T \widehat{\mathbf{W}}_e(\mathbf{m} - \mathbf{m}_{\mathrm{apr}}), \tag{8}$$

where $\widehat{\mathbf{W}}_e$ is a diagonal matrix:

$$\widehat{\mathbf{W}}_e = [(\widehat{\mathbf{m}} - \widehat{\mathbf{m}}_{\mathrm{apr}})^2 + e^2 \widehat{\mathbf{I}}]^{-1/2}. \tag{9}$$

This matrix depends on $\mathbf{m}$, that is why we call (8) a 'pseudo-quadratic' functional. We can introduce a linear transformation of the original model parameters into the space of the weighted model parameters:

$$\mathbf{m}^w = \widehat{\mathbf{W}}_e \mathbf{m}. \tag{10}$$

As a result of this transformation, we arrive at the traditional quadratic minimum norm functional, $s_{MN}(\mathbf{m}^w)$, for the weighted parameters $\mathbf{m}^w$:

$$s_{MS}(\mathbf{m}) = s_{MN}(\mathbf{m}^w) = (\mathbf{m}^w - \mathbf{m}^w_{\mathrm{apr}})^T (\mathbf{m}^w - \mathbf{m}^w_{\mathrm{apr}}) = \|\mathbf{m}^w - \mathbf{m}^w_{\mathrm{apr}}\|^2.$$

We also introduce the weighted data $\mathbf{d}^w$ as

$$\mathbf{d}^w = \widehat{\mathbf{W}}_d \mathbf{d}.$$

Using these notation, we can rewrite the parametric functional (2) as follows:

$$P^\alpha(\mathbf{m}^w) = \|\mathbf{A}^w(\mathbf{m}^w) - \mathbf{d}^w\|^2 + \alpha\|\mathbf{m}^w - \mathbf{m}^w_{\mathrm{apr}}\|^2$$

$$= (\mathbf{A}^w(\mathbf{m}^w) - \mathbf{d}^w)^*(\mathbf{A}^w(\mathbf{m}^w) - \mathbf{d}^w) + \alpha(\mathbf{m}^w - \mathbf{m}^w_{\mathrm{apr}})^T (\mathbf{m}^w - \mathbf{m}^w_{\mathrm{apr}}) \to \min, \tag{11}$$

where $\mathbf{A}^w(\mathbf{m}^w) = \widehat{\mathbf{W}}_d \mathbf{A}(\widehat{\mathbf{W}}_e^{-1} \mathbf{m}^w)$.

Note that the unknown parameters now are weighted model parameters, $\mathbf{m}^w$. In order to obtain the original model parameters, we have to apply inverse weighting to the result of minimization of the parametric functional (11):

$$\mathbf{m} = \widehat{\mathbf{W}}_e^{-1} \mathbf{m}^w. \tag{12}$$

Therefore, the problem of minimizing the parametric functional, given by equation (2), can be treated in a similar way to the minimization of the conventional Tikhonov functional (6). The only difference is that now we introduce some variable weighting matrix $\mathbf{W}_e$ for the model parameters which depends on the current model parameters. The minimization problem (11) can be solved using any gradient-type technique, say, by the conjugate gradient method (Zhdanov 2002).

Portniaguine and Zhdanov (1999) have developed a simplified approach to minimizing the parametric functional (11), using the re-weighted regularized conjugate gradient (RRCG) method. In the framework of this approach, the variable weighting matrix $\mathbf{W}_e$ is precomputed

on each iteration, $\mathbf{W}_e = \mathbf{W}_{en} = \mathbf{W}_e(\mathbf{m}_n)$, based on the values $\mathbf{m}_n$, obtained on the previous iteration. This linear transformation is updated after a fixed number of intermediate iterations, which is equivalent to the re-weighting of the model parameters. The advantage of this approach lies in its simplicity. The disadvantage is related to the fact that due to re-weighting, the misfit and stabilizing functionals can change, and even increase from iteration to iteration (Zhdanov 2002). In this paper we consider a different method of minimum support nonlinear inversion, which does not have the problems mentioned above and at the same time does not require minimization of a nonconvex functional. This technique is based on a new nonlinear parametrization, outlined below.

3. Minimum support nonlinear parametrization

We suggest using a nonlinear transformation instead of a set of variable linear transformations of the model parameters. In particular, we transform the nonquadratic minimum support stabilizing functional into a quadratic one by using a specially selected nonlinear transformation of the model parameters, described by the formula

$$\widetilde{m}_i = \frac{m_i - m_{\mathrm{apr}_i}}{\sqrt{\left(m_i - m_{\mathrm{apr}_i}\right)^2 + e^2}}, \tag{13}$$

where $\mathbf{m} = \{m_i\}, i = 1, \ldots, N_m$, is the original vector of the model parameters, and $\widetilde{\mathbf{m}} = \{\widetilde{m}_i\}, i = 1, \ldots, N_m$ is a new vector of the nonlinear parameters. The following inverse transform holds:

$$m_i - m_{\mathrm{apr}_i} = \frac{\widetilde{m}_i e}{\sqrt{1 - \widetilde{m}_i^2}}. \tag{14}$$

A transformation pair, (13) and (14), can be written in matrix notation as

$$\widetilde{\mathbf{m}} = \widehat{\mathbf{W}}_e(\mathbf{m} - \mathbf{m}_{\mathrm{apr}}) = [(\widehat{\mathbf{m}} - \widehat{\mathbf{m}}_{\mathrm{apr}})^2 + e^2 \widehat{\mathbf{I}}]^{-1/2}(\mathbf{m} - \mathbf{m}_{\mathrm{apr}}), \tag{15}$$

and

$$\mathbf{m} - \mathbf{m}_{\mathrm{apr}} = e[\widehat{\mathbf{I}} - \widehat{\widetilde{\mathbf{m}}}^2]^{-1/2}\widetilde{\mathbf{m}}, \tag{16}$$

where $\widehat{\widetilde{\mathbf{m}}}$ is an $N_m \times N_m$ diagonal matrix with the diagonal formed by nonlinear model parameters, $\widehat{\widetilde{\mathbf{m}}} = \mathrm{diag}(\widetilde{m}_1, \widetilde{m}_2, \ldots, \widetilde{m}_{N_m})$.

Substituting (16) into the Tikhonov parametric functional (2), we arrive at the conventional parametric functional with a minimum norm stabilizer

$$P[\widetilde{\mathbf{m}}] = \|\mathbf{A}^w(\mathbf{m}_{\mathrm{apr}} + e[\widehat{\mathbf{I}} - \widehat{\widetilde{\mathbf{m}}}^2]^{-1/2}\widetilde{\mathbf{m}}) - \mathbf{d}^w\|^2 + \alpha\|\widetilde{\mathbf{m}} - \widetilde{\mathbf{m}}_{\mathrm{apr}}\|^2. \tag{17}$$

We call the transformation, described by formulae (15) and (16), *a minimum support nonlinear parametrization* because it automatically ensures that the solution belongs to the class of models with minimum support.

Note that, in the implementation of the CG method, we have to calculate on every iteration n the Fréchet derivative matrix, $\widehat{\widetilde{\mathbf{F}}}_{(n)}$, with respect to a new model parameter vector $\widetilde{\mathbf{m}}_n$, which is equal to

$$\widehat{\widetilde{\mathbf{F}}}_{(n)} = e\widehat{\mathbf{F}}_{(n)}[\widehat{\mathbf{I}} - \widehat{\widetilde{\mathbf{m}}}_n^2]^{-3/2}, \tag{18}$$

where $\widehat{\mathbf{F}}_{(n)}$ is a Fréchet derivative of our original operator $\mathbf{A}^w$ at a point $\mathbf{m}_n$.

We should note, however, that in the case of a new minimum support nonlinear parametrization, we must consider a constrained optimization, with the absolute value of

$\widetilde{\mathbf{m}}$ being less than 1. In order to keep the true model parameter $\mathbf{m}$ within the known bounds, we also should impose additional constraints on $\widetilde{\mathbf{m}}$ and require that

$$\widetilde{m}_i^- < \widetilde{m}_i < \widetilde{m}_i^+, \tag{19}$$

where

$$\widetilde{m}_i^- = \frac{m_i^-}{\sqrt{m_i^{-2} + e^2}}, \qquad \text{and} \qquad \widetilde{m}_i^+ = \frac{m_i^+}{\sqrt{m_i^{+2} + e^2}}, \qquad \text{for} \quad i = 1, \ldots, N_m, \tag{20}$$

where $\mathbf{m}^+ = \{m_i^+\}$, and $\mathbf{m}^- = \{m_i^-\}$ are the known upper and lower boundaries for material property.

Therefore, we will use the gradient projection method (Bertsekas 1999), and take the projection of the parameters $\widehat{\mathbf{m}}$ onto the compact set, bounded by (19). We will perform minimization in a space of new parameters $\mathbf{m}^p \in (-\infty; \infty)$, and then take the projection onto this set, using the following function for the projection:

$$\widetilde{m}_i = \widetilde{m}_i^- + \frac{\widetilde{m}_i^+ + \widetilde{m}_i^-}{2} \left[\frac{2}{\pi} \tan^{-1} \left(\widetilde{m}_i^p - \frac{\widetilde{m}_i^+ + \widetilde{m}_i^-}{2} \right) + 1 \right].$$

This is a one-to-one transformation, which ensures (19) and has the inverse transformation:

$$\widetilde{m}_i^p = \frac{\widetilde{m}_i^+ + \widetilde{m}_i^-}{2} + \tan \frac{(\widetilde{m}_i - \widetilde{m}_i^-)\pi}{(\widetilde{m}_i^+ - \widetilde{m}_i) - \pi/2}. \tag{21}$$

The expression for the Fréchet derivative $\widehat{\widetilde{\mathbf{F}}}_{p(n)}$ with respect to a new model parameter vector $\mathbf{m}_n^p$ at the nth iteration of the CG method will take the form

$$\widehat{\widetilde{\mathbf{F}}}_{p(n)} = e\widehat{\mathbf{F}}_{(n)} \left[\mathbf{I} - \widehat{\widetilde{\mathbf{m}}}_n^2 \right]^{-3/2} \mathrm{diag} \left[\frac{\widetilde{m}_{ni}^+ - \widetilde{m}_{ni}^-}{\pi} \left(1 + \left(\widetilde{m}_{ni}^p \right)^2 \right)^{-1}, i = 1, \ldots, N_m \right]. \tag{22}$$

Finally, we can write the following regularized conjugate gradient (RCG) algorithm of the parametric functional minimization with two parametrizations:

$$\mathbf{r}_n = \mathbf{A}^w(\mathbf{m}_n) - \mathbf{d}^w, \tag{23}$$

$$\mathbf{l}_n = \mathrm{Re}[\widehat{\widetilde{\mathbf{F}}}_{p(n)}^* \mathbf{r}_n] + \alpha \widetilde{\mathbf{m}}_n^p, \tag{24}$$

$$\widetilde{\mathbf{l}}_n = \mathbf{l}_n + \widetilde{\mathbf{l}}_{n-1} \frac{\|\mathbf{l}_n\|}{\|\mathbf{l}_{n-1}\|}, \qquad \widetilde{\mathbf{l}}_0 = \mathbf{l}_0 \tag{25}$$

$$k_n = \widetilde{\mathbf{l}}_n^* \mathbf{l}_n [\|\widehat{\widetilde{\mathbf{F}}}_{p(n)} \widetilde{\mathbf{l}}_n \|^2 + \alpha \|\widetilde{\mathbf{l}}_n\|^2]^{-1}, \tag{26}$$

$$\widetilde{\mathbf{m}}_{n+1}^p = \widetilde{\mathbf{m}}_n^p - k_n \widetilde{\mathbf{l}}_n, \tag{27}$$

$$\widetilde{\mathbf{m}}_{n+1} = \widetilde{\mathbf{m}}^- + \frac{1}{2}(\widetilde{\mathbf{m}}^+ - \widetilde{\mathbf{m}}^-) \left(\frac{2}{\pi} \tan^{-1} \left(\widetilde{\mathbf{m}}_{n+1}^p \right) + 1 \right), \tag{28}$$

$$\mathbf{m}_{n+1} = e \left[\widehat{\mathbf{I}} - \widehat{\widetilde{\mathbf{m}}}_{n+1}^2 \right]^{-1/2} \widetilde{\mathbf{m}}_{n+1} + \mathbf{m}_{\mathrm{apr}}. \tag{29}$$

Note that algorithm (23)–(29) describes a conventional nonlinear conjugate gradient method applied to the Tikhonov parametric functional minimization (Zhdanov 2002, p 149). Vectors $\mathbf{l}_n$ are the gradient directions, while vectors $\widetilde{\mathbf{l}}_n$ represent the conjugate directions. In the inversion algorithm we use the real part of the Fréchet derivative $\widehat{\widetilde{\mathbf{F}}}_{p(n)}$, because the model parameters are described by real numbers, while the observed and predicted data can be complex numbers. The length of a step, k_n, is determined using a linear line search on every iteration.
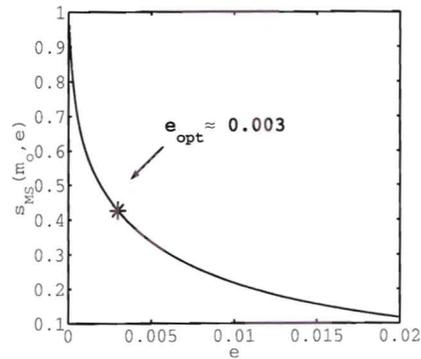
Figure 1. The values of the minimum support stabilizer for $\mathbf{m}_0$ versus parameter e.

The difficulty with nonlinear parameters is that, in the case of zero initial approximation ($\widetilde{\mathbf{m}} = 0$) in the conjugate gradient algorithm, we will get almost zeros for the Fréchet derivative, because $\mathbf{m}^p$ would be too large, and hence, we will not step away from the starting zero point. To overcome this difficulty, we use as a starting point for minimum support inversion a solution, $\mathbf{m}_0$, obtained with the minimum norm stabilizer.

Another difficulty is related to the problem of the optimum focusing parameter e calculation. Obviously, e should be chosen taking into account various aspects of the inverse problem. For example, e should not be too small, because it multiplies the Fréchet derivative (18). At the same time it should not be too large, because in this case the image will not be focused. In other words, the problem of selecting the focusing parameter e is very similar to the problem of choosing the regularization parameter α. Here we propose a method to estimate the optimum value of e, using the starting point $\mathbf{m}_0$ of the minimum support inversion.

Figure 1 presents a plot of the normalized value of the minimum support stabilizer versus the various values of e, computed for the given $\mathbf{m}_0$ according to the formula

$$s_{MS}(\mathbf{m}) = \mathbf{m}_0^T \left[\widehat{\mathbf{m}}_0^2 + e^2 \widehat{\mathbf{I}}\right]^{-1} \mathbf{m}_0.$$

It can be proved that this is a monotonically decreasing function, which goes to zero when e goes to $+\infty$, and to 1 when e goes to zero. Extremely small values of e may result in a singular behaviour of this functional (division by zero for zero elements of the diagonal matrix $\widehat{\mathbf{m}}_0$), while large values of e correspond to the standard minimum norm solution without focusing. That is why, by analogy with the L-curve method for the regularization parameter α (Hansen 1998), we chose e to be in the vicinity of the maximum convex curvature point. Our numerical studies show the effectiveness of this choice.

After introducing a trigonometric parametrization, we can proceed with the minimum support nonlinear parametrization and find the minimum of the Tikhonov parametric functional using the RCG algorithm (23)–(29). This completes the formulation of a new inversion scheme based on the minimum support nonlinear parametrization.

4. MT inversion

In this section, we will consider, as an illustration of the developed method, a solution of the geophysical magnetotelluric inverse problem. The MT method was introduced in geophysics

by Tikhonov (1950) and Cagniard (1953). It is based on measurements of the natural EM field at the surface of the earth. This field consists of a primary component of external origin and a secondary component that arises due to telluric currents induced in conductive regions of the earth by the primary field. The penetration depth of the primary field, and therefore of the telluric currents, increases with the period. The interpretation of magnetotelluric data is based on the calculation of the transfer functions between the horizontal components of the electric and magnetic fields, which form the impedance tensor $\hat{\mathbf{Z}}$ (Zhdanov and Keller 1994). The components of the impedance tensor depend primarily on the subsurface resistivity distribution over the penetration depth. Impedance measurements as a function of period can therefore be inverted for a resistivity model of the earth.

Following the traditional approach used in practical MT observations, we can calculate the apparent resistivities, ρ, and phases, ϕ, based on two off-diagonal elements of the MT tensor, Z_{xy} and Z_{yx}, at each observation point:

$$\rho_{xy} = \frac{1}{\omega\mu_0}|Z_{xy}|^2, \qquad \rho_{yx} = \frac{1}{\omega\mu_0}|Z_{yx}|^2, \tag{30}$$

$$\phi_{xy} = \tan^{-1}\frac{\operatorname{Im} Z_{xy}}{\operatorname{Re} Z_{xy}}, \qquad \phi_{yx} = \tan^{-1}\frac{\operatorname{Im} Z_{yx}}{\operatorname{Re} Z_{yx}}, \tag{31}$$

where the quantities ρ_{xy} and ϕ_{xy} are assigned to the nominal transverse magnetic (TM) mode, whereas ρ_{yx} and ϕ_{yx} are assigned to the nominal transverse electric (TE) mode. Note that this 2D nomenclature is artificial and approximate in nature for 3D structures. However, it is used in practical MT observations (Wannamaker 1997, Zhdanov *et al* 2000a). Note also that, in the inversion algorithm, we actually use the logarithm of apparent resistivity and phase in radians.

Therefore, the solution of the inverse problem requires numerical modelling of the apparent resistivities and phases in each step of the iteration process. The components of the impedance tensor are determined from the horizontal components of the electric and magnetic fields in every observation point. The corresponding technique of solving this problem is outlined, for example, in Zhdanov and Keller (1994) and Berdichevsky and Dmitriev (2002). This procedure is extremely time consuming, which results in enormous calculations for the solution of the inverse problem. In order to overcome these computational difficulties, following Zhdanov and Hursán (2000) we suggest using an approximate solution based on the quasi-analytical (QA) approximation on the initial stage of the iterative inversion. A description of the basic principles of the QA approximation can be found in Zhdanov *et al* (2000b) and Zhdanov (2002, pp 248–51).

The approximate QA forward operators can be used for computing the components of the impedance tensor components in (30) and (31). These operators speed up significantly the calculations on each step of the inversion. In a general case, the corresponding formulae can be expressed by an operator equation including the data vector **d** and the vector of model parameters **m** as

$$\mathbf{d} = \mathbf{A}(\mathbf{m}), \tag{32}$$

where **A** is the forward operator symbolizing the governing equations of the MT impedance modelling problem, **m** is the vector of the unknown conductivity distribution (model parameters) and **d** is the vector formed by observed values of the apparent resistivities and phases at the observation points. We can apply now the RCG algorithm (23)–(29) with the minimum support nonlinear parametrization to solve the MT inverse problem (32). Note that the computation of the Fréchet derivative matrix, required by this algorithm, can be made on the basis of QA approximation as well (Zhdanov and Hursán 2000).

Application of the QA approximation to forward modelling and Fréchet derivative computations speeds up the calculation dramatically. However, in order to control the accuracy of the inversion, our method allows application of rigorous forward modelling in the final steps of the inversion procedure. We use an integral equation forward modelling code based on the contraction integral equation method, which improves the convergence rate of the iterative solvers (Hursán and Zhdanov 2002). A few last iterations with rigorous forward modelling require much more time than all previous iterations with the QA approximation. However, application of the rigorous solver improves the resolution of the inverse method and helps to generate a more correct image of the target.

5. Numerical modelling results

The 3D MT inversion algorithm described above and the corresponding computer code have been carefully tested on synthetic models. We present below some numerical examples of the MT data inversion with this new method.

5.1. A conductive dike model

Consider a homogeneous half-space with a background resistivity of $\rho_b = 100 \, \Omega$m, containing a conductive dike. The resistivity of the inhomogeneity is 3 Ωm. The top of the dike is at a depth of 200 m, and its bottom is at a depth of 600 m beneath the surface. This model is excited by a plane EM wave source. The x and y components of the anomalous magnetic and electric fields for four different frequencies (1, 10, 100 and 1000 Hz) have been simulated at 225 receiver points arranged on a homogeneous grid, using the integral equation forward modelling code INTEM3D (Hursán and Zhdanov 2002). The coordinates of the receiver grid are the following: x and y from -700 to 700 every 100 m. The receiver system is located at the surface of the earth. The EM field components were recalculated into MT apparent resistivity and phase, using the standard formulae (Berdichevsky and Dmitriev 2002). The area of inversion is covered by a homogeneous mesh consisting of $16 \times 25 \times 8$ cubic cells surrounding the anomalous structure to be inverted. Each cell has a dimension of 100 m in the x, y and z directions. We select a focusing parameter based on the corresponding L-curve analysis, described in section 3, as follows: $e = 0.016$. The upper and lower bounds of the total resistivity obtained from the inversion are set within an interval of 1–100 Ωm.

Figure 2 shows the true model (panel (a)), discretization grid and MT station positions for this model (panel (b)), inversion results after 15 iterations with the minimum norm stabilizer (panel (c)), intermediate result after 60 iterations with the minimum support stabilizer and QA forward modelling (panel (d)) and the final result after 5 additional iterations of the RCG method with minimum support inversion and rigorous full forward modelling (panel (e)). In figure 2, we present three-dimensional images of the resistivity distribution with the volume rendering. The cut-off level of the resistivity for these images is shown in the corresponding panels. For example, the cut-off level $\rho < 20 \, \Omega$m means that only the cells with a value of resistivity less than 20 Ωm are displayed. Figure 3 presents the inversion curves, parametric functional, $P[\alpha]$, stabilizer, $S[m]$, misfit, $\phi[m]$, and elapsed time versus iteration number.

Note that, at iteration number 16, we observe a sudden decrease in $S[m]$ because at this iteration the code switches from the minimum norm to the minimum support stabilizer. The transition at the 76th iteration from the approximate forward modelling operator to the full integral equation solver results in a notable increase in the misfit and parametric functional at this iteration. Also, due to the high conductivity contrast, the predicted conductivity is slightly smaller than the true conductivity. Nevertheless, the inversion with the minimum
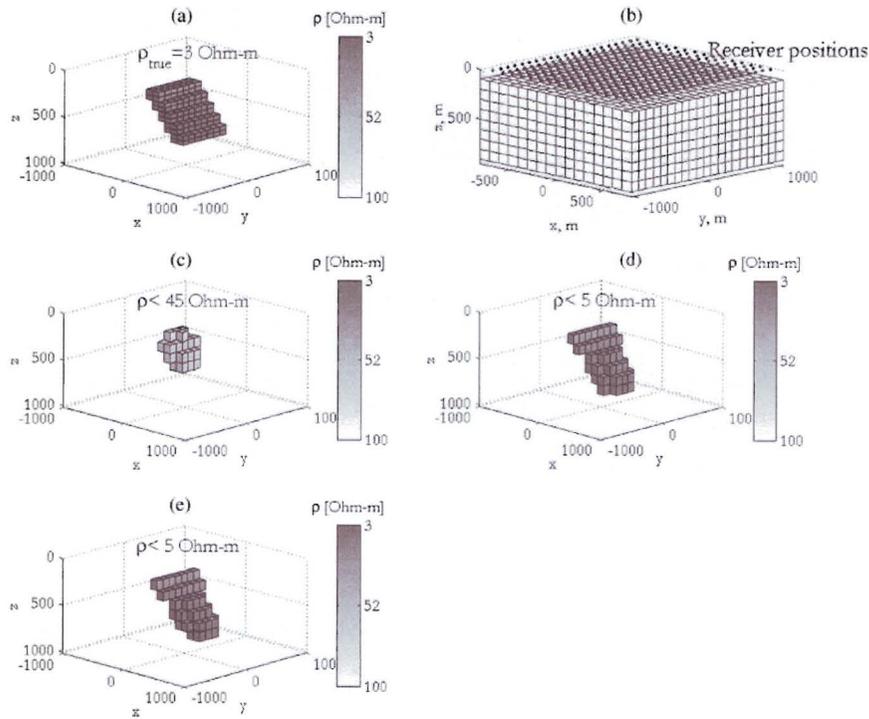
Figure 2. Conductive dike model. (a) True model; (b) discretization grid and MT station positions; (c) inversion results after 15 iterations with the minimum norm stabilizer; (d) intermediate result after 60 iterations with the minimum support stabilizer and QA forward modelling; (e) final result after 5 iterations of the RCG method with the minimum support inversion and rigorous full forward modelling.

support stabilizer helps us to obtain an image with much greater contrast, than minimum norm inversion, as one can see from figure 2, where the cut-off value for the image in panel (c) is four times greater than for panels (d) and (e). The shape and position of the recovered body are predicted quite well.

For comparison, we present in figure 4 the vertical cross-sections of the true model (panel (a)), the inversion result with the minimum norm stabilizer (panel (b)), the intermediate result with the minimum support stabilizer and QA forward modelling (panel (c)) and the final sharp inversion result (panel (d)). One can see that the minimum norm result underestimates the true conductivity, while the sharp inversion reconstructs an image very close to the true model.

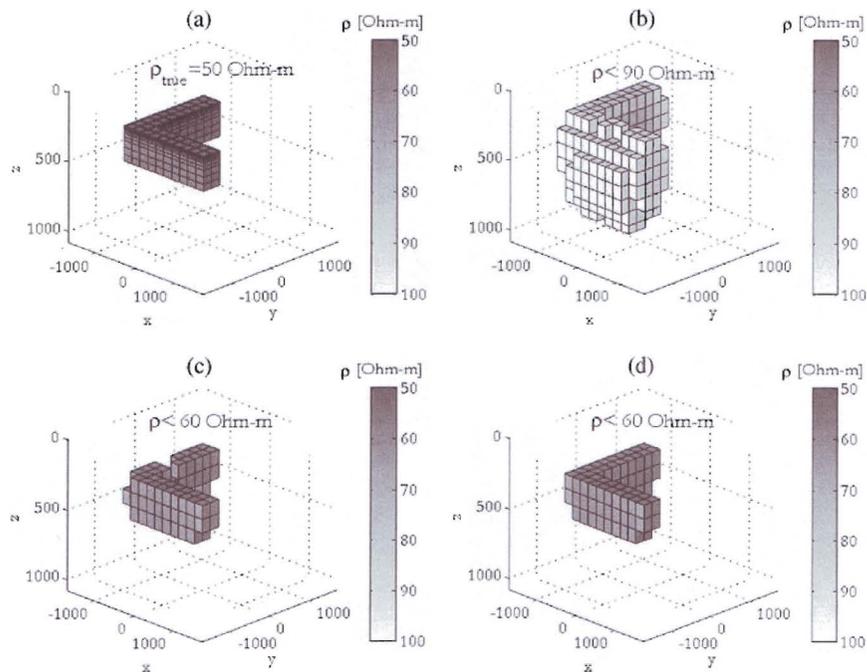### 5.2. Model of an L-shaped conductive structure

The next model we considered was an L-shaped conductive structure. The top of the structure is 250 m below the earth's surface, and the bottom is 500 m below. The resistivity of the anomalous zone is 50 Ωm, and it was embedded in a 100 Ωm homogeneous half-space. The model is excited by the plane wave source (0.1, 1, 10 and 100 Hz). Receivers are placed every 200 m along the x and y directions from -1600 to 1600 m. We inverted the synthetic MT data
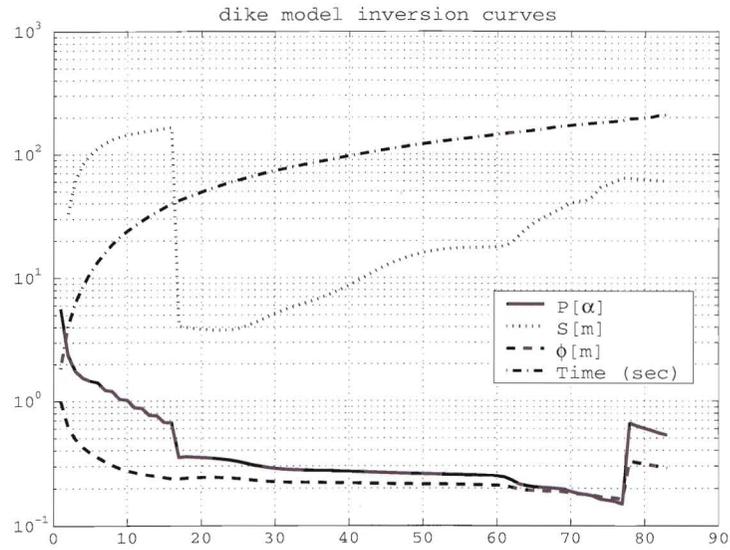
Figure 3. Inversion curves for the dike model: parametric functional, $P[\alpha]$, stabilizer, $S[m]$, misfit, $\phi[m]$, and elapsed time versus iteration number.
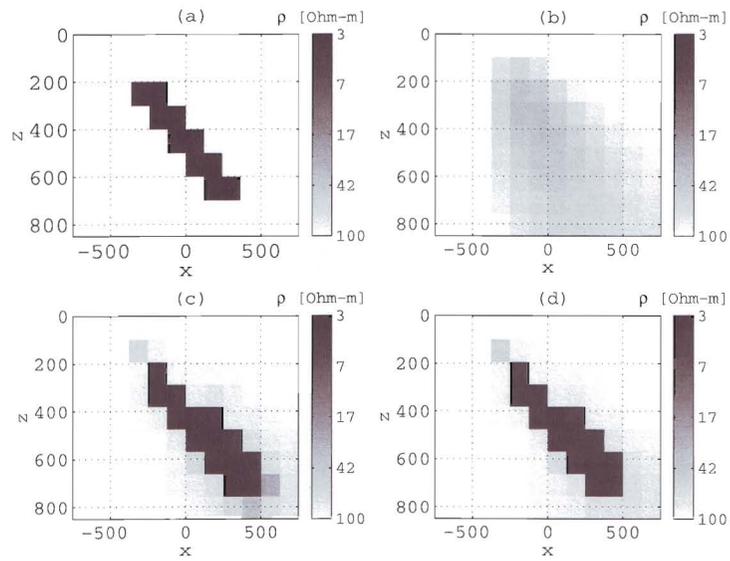
Figure 4. Dike model: vertical cross-sections of the true model (panel (a)), the inversion result with the minimum norm stabilizer (panel (b)), the intermediate result with the minimum support stabilizer and QA forward modelling (panel (c)) and the final sharp inversion result (panel (d)).

Figure 5. L-shaped conductive structure: (a) true model; (b) inversion results after 20 iterations with the minimum norm stabilizer; (c) intermediate result after 63 iterations with the minimum support stabilizer and QA forward modelling; and (d) final result after 5 iterations of the RCG method with the minimum support inversion and rigorous full forward modelling.

computed for this model using the full integral equation solver at the final stage of inversion. The area of inversion was covered by a rectangular grid consisting of $18 \times 18 \times 8$ cells, and every cell was $200 \times 200 \times 100$ m^3. The selected focusing parameter is equal to $e = 0.004$. The upper and lower bounds of the total resistivity obtained from the inversion are set within an interval of 10–100 Ωm.
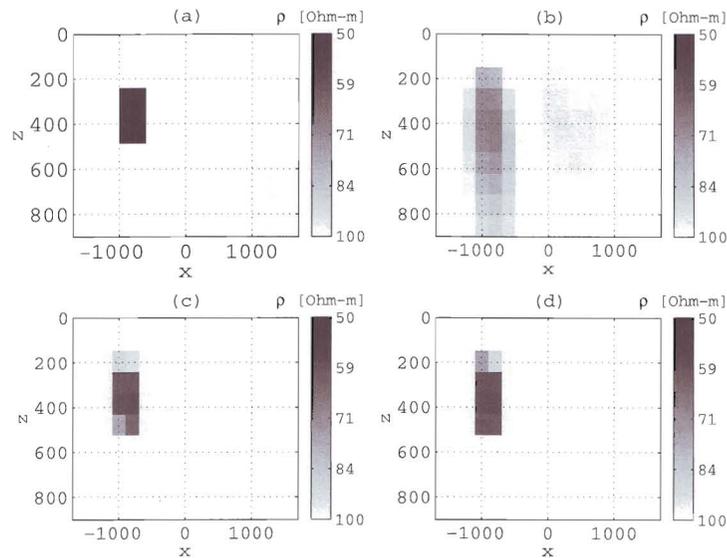
Figure 5 presents the true model (panel (a)), inversion results after 20 iterations with the minimum norm stabilizer (panel (b)), an intermediate result after 63 iterations with the minimum support stabilizer and QA forward modelling (panel (c)) and the final result after 5 additional iterations of the RCG method with the minimum support inversion and rigorous forward modelling (panel (d)). Figure 6 shows the inversion curves, parametric functional, $P[\alpha]$, stabilizer, $S[m]$, misfit, $\phi[m]$, and elapsed time versus the iteration number. Figure 7 shows the vertical cross-sections of the true model (panel (a)), the inversion result with the minimum norm stabilizer (panel (b)), the intermediate result with the minimum support stabilizer and QA forward modelling (panel (c)) and the final sharp inversion result (panel (d)). The recovered image reconstructs well the original model.

5.3. 'Open box' conductive structure

We present a similar result for the 'open box' model, simulating a conductive syncline structure. The resistivity of the anomalous zone is 50 Ωm, while for the surrounding medium it is equal
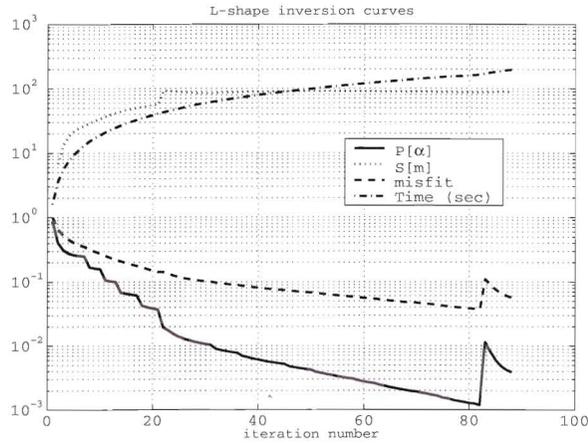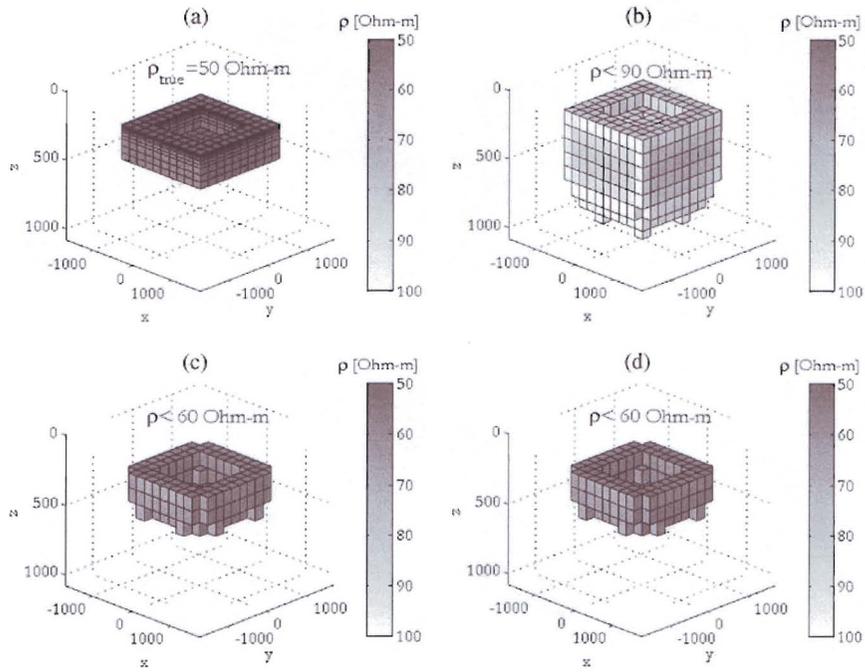
Figure 6. Inversion curves for the L-shaped model, parametric functional, $P[\alpha]$, stabilizer, $S[m]$, misfit, $\phi[m]$, and elapsed time versus iteration number.
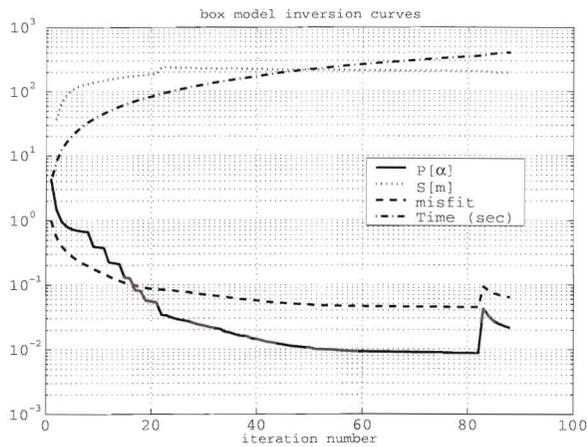

Figure 7. L-shaped model: vertical cross-sections of the true model (panel (a)), the inversion result with the minimum norm stabilizer (panel (b)), the intermediate result with the minimum support stabilizer and QA forward modelling (panel (c)) and the final sharp inversion result (panel (d)).

to 100 Ωm. The model is excited by the plane wave source (0.1, 1, 10 and 100 Hz). We use the same MT observation system, the same grid and the same parameters for inversion as for a model shown in figure 5.

Figure 8 presents the true model (panel (a)), inversion results after 20 iterations with the minimum norm stabilizer (panel (b)), an intermediate result after 63 iterations with the

Figure 8. 'Open box' conductive structure: (a) true model; (b) inversion results after 20 iterations with the minimum norm stabilizer; (c) intermediate result after 50 iterations with the minimum support stabilizer and QA forward modelling and (d) final result after 5 iterations of the RCG method with minimum support inversion and rigorous full forward modelling.

Figure 9. 'Open box model': inversion curves, parametric functional, $P[\alpha]$, stabilizer, $S[m]$, misfit, $\phi[m]$, and elapsed time versus iteration number.
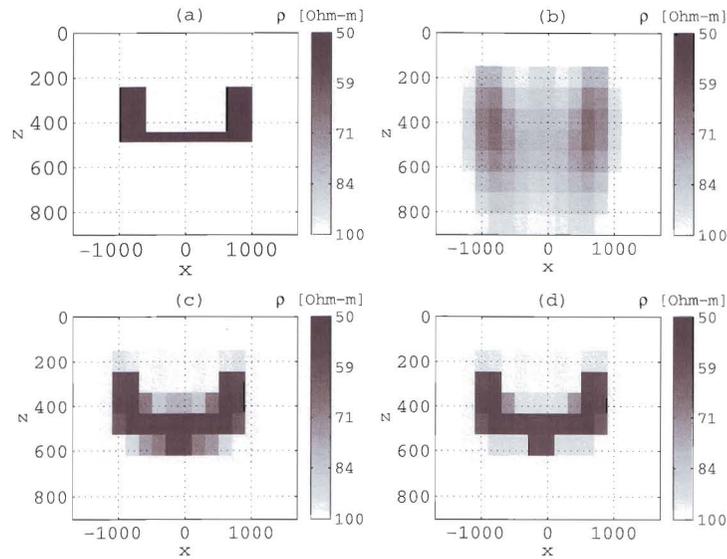
Figure 10. 'Open box' model: vertical cross-sections of the true model (panel (a)), the inversion result with the minimum norm stabilizer (panel (b)), the intermediate result with the minimum support stabilizer and QA forward modelling (panel (c)), and the final sharp inversion result (panel (d)).

minimum support stabilizer and QA forward modelling (panel (c)) and the final result after 5 additional iterations of the RCG method with the minimum support inversion and rigorous forward modelling (panel (d)). Figure 9 shows the inversion curves, parametric functional, $P[\alpha]$, stabilizer, $S[m]$, misfit, $\phi[m]$, and elapsed time versus iteration number. Figure 10 shows the vertical cross-sections of the true model (panel (a)), the inversion result with the minimum norm stabilizer (panel (b)), an intermediate result with the minimum support stabilizer and QA forward modelling (panel (c)) and the final sharp inversion result (panel (d)). The recovered image reconstructs well the original model. One can observe, however, that some artefacts appeared beneath the inhomogeneity. These artefacts can be explained by a simple fact that the sensitivity of the data to the bottom of the model is smaller than the sensitivity to the top part of the conductive body.

6. Conclusion

We have demonstrated in this paper that regularized inversion with a minimum support stabilizer can be implemented by using a specially designed nonlinear parametrization of the model parameters. This parametrization plays the same role as the transformation into the space of the weighted model parameters, introduced in the original papers on focusing inversion. It allows transforming the nonquadratic minimum support stabilizer into the traditional quadratic minimum norm stabilizer, which simplifies the solution of the inverse problem. Also, using this transformation and gradient projection method, we impose the upper and lower bounds on the model parameter distribution, and automatically ensure that the solution belongs to the class of models with the minimum support.

Based on this new parametrization, we developed a new algorithm of the rapid MT inversion and investigated the effectiveness of the new approach on synthetic models. The model study shows a good performance of the method with the new nonlinear parametrization.

Acknowledgments

References

Berdichevsky M N and Dmitriev V I 2002 Magnetotellurics in the context of the theory of ill-posed problems (Tulsa, OK: Society of Exploration Geophysicists) p 215

Bertsekas D P 1999 *Nonlinear Programming* 2nd edn (Belmont, MA: Athena Scientific) p 791

Cagniard L 1953 Theory of magneto-telluric method of geophysical prospecting *Geophysics* **18** 605–35

Geman D and Reynolds G 1992 Constrained restoration and the recovery of discontinuities *IEEE Trans. Pattern Anal. Mach. Intell.* **14** 367–83

Geman D and Yang C 1995 Nonlinear image recovery with half-quadratic regularization *IEEE Trans. Image Process.* **4** 932–46

Hansen C 1998 Rank-deficient and discrete ill-posed problems *Numerical Aspects of Linear Inversion (SIAM Monographs on Mathematical Modeling and Computation)* (Philadelphia, PA: SIAM) p 247

Hursán G and Zhdanov M S 2002 Contraction integral equation method in 3-D electromagnetic modeling *Radio Sci.* **37** 1089

Last B J and Kubik K 1983 Compact gravity inversion *Geophysics* **48** 713–21

Lobel P, Blanc-Feraud L, Pichot Ch and Barlaud M 1997 A new regularization scheme for inverse scattering *Inverse Problems* **13** 403–10

Mehanee S and Zhdanov M S 2002 Two-dimensional magnetotelluric inversion of blocky geoelectrical structures *J. Geophys. Res.* **107** 10.1029

Portniaguine O and Zhdanov M S 1999 Focusing geophysical inversion images *Geophysics* **64** 874–87

Portniaguine O and Zhdanov M S 2002 3-D magnetic inversion with data compression and image focusing *Geophysics* **67** 1532–41

Tikhonov A N 1950 On determining electric characteristics of the deep layers of the earth's crust *Dokl. Akad. Nauk SSSR* **73** 295–7

Tikhonov A N and Arsenin V Y 1977 *Solution of Ill-Posed Problems* (Washington, DC: V H Winston and Sons) p 258

Vogel C R 1997 Nonsmooth regularization *Inverse Problems in Geophysical Applications* ed H W Engl, A K Louis and W Rundell (Philadelphia, PA: SIAM) pp 1–11

Wannamaker P 1997 Tensor CSAMT survey over the Sulphur Springs thermal area, Valles Caldera, New Mexico, USA: Part II. Implications for CSAMT methodology *Geophysics* **62** 466–76

Zhdanov M S 2002 *Geophysical Inverse Theory and Regularization Problems* (Amsterdam: Elsevier) p 628

Zhdanov M S, Dmitriev V I, Fang S and Hursán G 2000b Quasi-analytical approximations and series in electromagnetic modeling *Geophysics* **65** 1746–57

Zhdanov M S, Fang S and Hursan G 2000a Electromagnetic inversion using quasi-linear approximation *Geophysics* **65** 1501–13

Zhdanov M S and Hursán G 2000 3-D electromagnetic inversion based on quasi-analytical approximation *Inverse Problems* **16** 1297–322

Zhdanov M S and Keller G W 1994 *The Geoelectrical Methods in Geophysical Exploration* (Amsterdam: Elsevier) p 873