

Resistivity imaging with controlled-source electromagnetic data: depth and data weighting

R.-É. Plessix and W.A. Mulder

Shell International E&P, Kessler Park 1, 2288 GS Rijswijk, The Netherlands

E-mail: reneedouard.plessix@shell.com, wim.mulder@shell.com

Abstract. We discuss some computational aspects of resistivity imaging by inversion of offshore controlled-source electromagnetic data. We adopt the classic approach to imaging by formulating it as an inverse problem. A weighted least-squares functional measures the misfit between synthetic and observed data. Its minimization by a quasi-Newton algorithm requires the gradient of the functional with respect to the model parameters. We compute the gradient with the adjoint-state technique. Preconditioners can improve the convergence of the inversion. Diagonal preconditioner based on a Born approximation are commonly used. In the context of CSEM inversion, the Born approximation is not really accurate, this limits the possibility of estimating a correct approximation of the Hessian in a smooth medium or, in fact, in any reference background that does not roughly account for the resistors. We hence rely on the limited memory BFGS approximation of the inverse of the Hessian and we improve the inversion convergence with the help of a heuristic data and depth weighting. Based on a numerical example, we show that a simple exponential depth weighting combined with an offset or frequency data weighting significantly improves the convergence rate of a deep-water controlled-source electromagnetic data inversion.

AMS classification scheme numbers: 86A22,86A25,35R30

Submitted to: *Inverse Problems*

1. Introduction

A classic offshore controlled-source electromagnetic (CSEM) survey consists of towing a horizontal electric dipole close to the sea floor and recording the electric and magnetic response on stations laid down on the sea floor. CSEM data are used in the oil and gas industry to detect hydrocarbons because they are sensitive to thin resistive horizontal layers [1, 2, 3]. The physical explanation is that the highly resistive layers strongly modify the current lines emitted by the horizontal electric dipole [4]. In the presence of resistors in the earth, the electromagnetic waves diffuse in the resistor at a higher “velocity” than in the surrounding conductive sediments and radiate back to the surface. These physics cannot be modeled by a Born (single scatterer) approximation in a smooth (at the skin depth scale) resistivity model [5]. This impacts the imaging of CSEM data [6].

In this paper, we discuss some computational and physical aspects of resistivity imaging from CSEM data. We formulate the imaging problem as an inverse problem [7, 8, 9]. Given electric and magnetic data, $\mathbf{e}_{s,r,f}^{obs}$ and $\mathbf{h}_{s,r,f}^{obs}$, generated by electric emitting dipoles at positions \mathbf{x}_s and frequencies f and observed at the receiver positions \mathbf{x}_r , the inverse problem consists of finding a conductivity that minimizes the weighted least-squares functional

$$J(\mathbf{p}) = \frac{1}{2} \sum_{\omega} \sum_s \sum_r \left\{ \|W_{s,r,f}^e(\mathbf{e}_{s,r,f}[\sigma(\mathbf{p})] - \mathbf{e}_{s,r,f}^{obs})\|^2 + \|W_{s,r,f}^h(\mathbf{h}_{s,r,f}[\sigma(\mathbf{p})] - \mathbf{h}_{s,r,f}^{obs})\|^2 \right\} + R(\mathbf{p}). \quad (1)$$

Here, $\mathbf{e}_{s,r,f}$ and $\mathbf{h}_{s,r,f}$ represent the computed response for a given conductivity σ . $W_{s,r,f}^e$ and $W_{s,r,f}^h$ are suitable data weights. The subscripts s , r , and f mark the dependency on source, receiver, and frequency, respectively. The conductivity σ depends on the model parameters \mathbf{p} , a vector in \mathbb{R}^M . In 3D, the number of model parameters, namely unknowns, is typically in the range of 1 to 20 million. Finally, R is a regularization term.

In complex media, the electric and magnetic response is computed by numerically solving Maxwell’s equations. In the earth, the conduction currents satisfy Ohm’s constitutive law and the electromagnetic waves are diffusive [10]. Different methods can be used to discretize and solve those equations, for instance [11, 12, 13, 14, 15, 16, 17] most of them based on the staggered scheme proposed by Yee [18]. In this work, we discretized the Maxwell/Ohm equations with the Finite Integration Technique (FIT) [17] and solved the linear system with a preconditioned Krylov method, BI-CGSTAB2, [19, 20]. The preconditioner for the linear iterative solver is a single multigrid cycle [21, 22]. Combined with an automatic frequency dependent regridding [23], this leads to an efficient solver for CSEM modeling.

Although the 3D numerical solvers are efficient, the computational cost of stochastic methods for the minimization of the least-squares functional (1) are still too high. We therefore solve the inverse problem by gradient optimization. Several authors have used this approach to invert magnetotelluric data, crosswell data, or CSEM data, for

instance [2, 14, 24, 25, 26, 27, 28, 29]. The least-squares functional can be inverted by a non-linear conjugate gradient optimization [25, 27], a Gauss-Newton optimization [30], or a quasi-Newton optimization [30, 31]. We applied a limited memory quasi-Newton approach, more precisely a limited memory BFGS (Broyden, Fletcher, Goldfarb and Shanno) method [32, 33] that is less commonly used in geophysics. The update \mathbf{p}^k at the iteration k is given by

$$\mathbf{p}^k = \mathbf{p}^{k-1} - \alpha^{k-1} \mathbf{K}^{k-1} \nabla_{\mathbf{p}} J(\mathbf{p}^{k-1}), \quad (2)$$

where α^{k-1} is a real scalar obtained by a suitable line-search algorithm and \mathbf{K}^{k-1} is the approximated inverse of the Hessian given by the BFGS method. This inversion can be preconditioned to improve convergence, for instance see [25, 30]. A common approach is to use the inverse of the diagonal of the Hessian, generally computed in a smooth initial model. The difficulty with CSEM inversion is that the Hessian is not necessarily diagonal dominant, hence this approach may be of limited relevance. Another classic approach is the Gauss-Newton approximation of the Hessian [30]. This requires the computation of the Fréchet derivatives, which can be expensive except in simplified resistivity models. Also, its relevance may depend on the accuracy of the Born (linearized) approximation. Our formulation follows the classic inverse problem formulation where the gradient is computed with the adjoint method [7, 9]. Newman and Hoversten [26] and Newman and Boggs [30] have for instance described this method in the context of controlled source electromagnetic surveys and Mackie and Radden [25] in the context of magnetotelluric surveys. In this paper, we decided to formulate the preconditioning as a scaling of the model parameters, which amounts to a depth weighting, and to use the BFGS method to approximate the inverse of the Hessian. Our approach differs from the approach of Newman and Boggs [30] in the preconditioner and in the data weighting. We do not base our preconditioner on an approximation of the Hessian in a reference resistivity model due to the limited accuracy of Born approximation for the low frequency electromagnetic applications in geophysics [5].

The outline of the paper is as follows. We first describe the gradient calculation of the discretized problem with the FIT using the adjoint-state method [35]. While this part is not new, it is important to notice the implementation advantage of this approach compared to the classic finite-difference scheme on stretched grids. Secondly, we discuss the scaling for CSEM applications. Then, we present synthetic numerical results to show the relevance of the depth weighting approach in combination with some commonly used data weighting schemes and we display an inversion result on the Troll real data. Finally, we summarize our conclusions.

2. Discrete formulation

The problem is formulated in the frequency domain because we invert the data at only a few frequencies. We model the CSEM response by solving Maxwell's equations. The

conduction currents satisfy Ohm's constitutive law in a conductive medium (see [10] for details). In the frequency domain, the second-order electromagnetic equation reads

$$\nu\mu_0\omega\tilde{\sigma}(\mathbf{x},\omega)\mathbf{E}_s(\mathbf{x},\omega) - \nabla \times \mu_r^{-1}(\mathbf{x})\nabla \times \mathbf{E}_s(\mathbf{x},\omega) = -\mathbf{F}_s(\mathbf{x},\omega), \quad (3)$$

with a source term

$$\mathbf{F}_s(\mathbf{x}_s,\omega) = \omega\mu_0\mathbf{J}_s(\mathbf{x},\omega) + \nabla \times \mu_r^{-1}(\mathbf{x})\mathbf{M}_s(\mathbf{x}_s,\omega).$$

\mathbf{E}_s is the electric field, \mathbf{M}_s the magnetic source, \mathbf{J}_s the electric source, and $\tilde{\sigma}(\mathbf{x},\omega) = \sigma(\mathbf{x}) - \omega\varepsilon_0\varepsilon_r(\mathbf{x})$ with conductivity σ , the inverse of the electric resistivity. The vacuum dielectric permittivity is denoted by ε_0 , ε_r is the relative dielectric permittivity, μ_0 is the vacuum magnetic permeability, μ_r the relative magnetic permeability, and $\omega = 2\pi f$ the angular frequency. Finally, $\mathbf{x} \in \Omega$ is a subsurface point inside a domain $\Omega \subset \mathbb{R}^3$. We adopted the Fourier convention $f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega)e^{-i\omega t}d\omega$. On the boundary Γ of Ω , we impose the perfectly electrical conductor boundary conditions

$$\mathbf{n}(\mathbf{x}) \times \mathbf{E}_s(\mathbf{x},\omega) = 0, \quad (4)$$

with \mathbf{n} the outward normal of Γ . We obtain the electric data, \mathbf{e} , and the magnetic data, \mathbf{h} , by sampling the electric field at the receiver positions \mathbf{x}_r :

$$\begin{cases} \mathbf{e}_{s,r,f}[\sigma] = \mathbf{S}^{ee}(\mathbf{x}_r, \mathbf{x})\mathbf{E}_s(\mathbf{x},\omega); \\ \mathbf{h}_{s,r,f}[\sigma] = \mathbf{S}^{he}(\mathbf{x}_r, \mathbf{x})\mathbf{E}_s(\mathbf{x},\omega). \end{cases} \quad (5)$$

\mathbf{S}^{ee} and \mathbf{S}^{he} are two sampling operators that map the fields to the receiver position \mathbf{x}_r . \mathbf{S}^{he} also takes the relation $\mathbf{H}_s = \frac{1}{\omega\mu_0\mu_r}\nabla \times \mathbf{E}_s$ into account.

Because electromagnetic waves diffuse in a conductive medium, which means that their amplitudes decrease exponentially, we often discretize the electromagnetic equations on a stretched grid to save computer time. The domain Ω is paved with a tensor-product cartesian grid $\mathbf{x}_g \otimes \mathbf{y}_g \otimes \mathbf{z}_g$, with $\mathbf{x}_g = (x_k)$ (resp. $\mathbf{y}_g = (y_l)$ and $\mathbf{z}_g = (z_m)$) an array of positions in the x -direction (resp. y and z -direction) with variable spacing. The electromagnetic equations are classically discretized by primal and dual grids. The primal grid is defined by the nodes (x_k, y_l, z_m) and the dual grid by the nodes $(x_{k+\frac{1}{2}}, y_{l+\frac{1}{2}}, z_{m+\frac{1}{2}})$. Here, $x_{k+\frac{1}{2}} = \frac{1}{2}(x_k + x_{k+1})$, $y_{l+\frac{1}{2}} = \frac{1}{2}(y_l + y_{l+1})$, and $z_{m+\frac{1}{2}} = \frac{1}{2}(z_m + z_{m+1})$. The spacing on the primal or dual grid is denoted by $h_\alpha^x = x_{\alpha+\frac{1}{2}} - x_{\alpha-\frac{1}{2}}$, $h_\beta^y = y_{\beta+\frac{1}{2}} - y_{\beta-\frac{1}{2}}$, and $h_\gamma^z = z_{\gamma+\frac{1}{2}} - z_{\gamma-\frac{1}{2}}$ with $\alpha = k$ or $\alpha = k + \frac{1}{2}$, $\beta = l$ or $\beta = l + \frac{1}{2}$, and $\gamma = m$ or $\gamma = m + \frac{1}{2}$. The conductivity, the relative magnetic permeability, and the relative dielectric permittivity are discretized as averages per cell. We denote their values by $\sigma_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}}$, $\mu_{r,k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}}^{-1}$, and $\varepsilon_{r,k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}}$. These can be interpreted as values at the center of each cell.

One of the classic discretizations of equation (3) is the centered finite difference method (FDM) on a staggered grid. On a regular spacing grid, this is known as Yee's scheme [18]. An extension of the scheme to stretched grids can be obtained with the FIT [17]. This scheme differs from the one directly obtained from finite difference by

some volume factors (see Appendix A). The FIT, based on an integral representation of the electromagnetic equations and flux quantities, leads to the following discretization:

$$\begin{aligned}
& \frac{U_{x,k,l+\frac{1}{2},m+\frac{1}{2}}}{\mathcal{M}_{k,l+\frac{1}{2},m+\frac{1}{2}}} - \frac{\mathbf{E}_{z,k,l+1,m+\frac{1}{2}} - \mathbf{E}_{z,k,l,m+\frac{1}{2}}}{h_{l+\frac{1}{2}}^y} + \frac{\mathbf{E}_{y,k,l+\frac{1}{2},m+1} - \mathbf{E}_{y,k,l+\frac{1}{2},m}}{h_{m+\frac{1}{2}}^z} = 0; \\
& \frac{U_{y,k+\frac{1}{2},l,m+\frac{1}{2}}}{\mathcal{M}_{k+\frac{1}{2},l,m+\frac{1}{2}}} - \frac{\mathbf{E}_{x,k+\frac{1}{2},l,m+1} - \mathbf{E}_{x,k+\frac{1}{2},l,m}}{h_{m+\frac{1}{2}}^z} + \frac{\mathbf{E}_{z,k+1,l,m+\frac{1}{2}} - \mathbf{E}_{z,k,l,m+\frac{1}{2}}}{h_{k+\frac{1}{2}}^x} = 0; \\
& \frac{U_{z,k+\frac{1}{2},l+\frac{1}{2},m}}{\mathcal{M}_{k+\frac{1}{2},l+\frac{1}{2},m}} - \frac{\mathbf{E}_{y,k+1,l+\frac{1}{2},m} - \mathbf{E}_{y,k,l+\frac{1}{2},m}}{h_{k+\frac{1}{2}}^x} + \frac{\mathbf{E}_{x,k+\frac{1}{2},l+1,m} - \mathbf{E}_{x,k+\frac{1}{2},l,m}}{h_{l+\frac{1}{2}}^y} = 0; \\
& \mathcal{S}_{k+\frac{1}{2},l,m} \mathbf{E}_{x,k+\frac{1}{2},l,m} - \left(\frac{U_{z,k+\frac{1}{2},l+\frac{1}{2},m}}{h_{l+\frac{1}{2}}^y} - \frac{U_{z,k+\frac{1}{2},l-\frac{1}{2},m}}{h_{l-\frac{1}{2}}^y} \right) + \left(\frac{U_{y,k+\frac{1}{2},l,m+\frac{1}{2}}}{h_{m+\frac{1}{2}}^z} - \frac{U_{y,k+\frac{1}{2},l,m-\frac{1}{2}}}{h_{m-\frac{1}{2}}^z} \right) = \\
& \quad -V_{k+\frac{1}{2},l,m} \mathbf{F}_{x,k+\frac{1}{2},l,m}; \\
& \mathcal{S}_{k,l+\frac{1}{2},m} \mathbf{E}_{y,k,l+\frac{1}{2},m} - \left(\frac{U_{x,k,l+\frac{1}{2},m+\frac{1}{2}}}{h_{m+\frac{1}{2}}^z} - \frac{U_{x,k,l+\frac{1}{2},m-\frac{1}{2}}}{h_{m-\frac{1}{2}}^z} \right) + \left(\frac{U_{z,k+\frac{1}{2},l+\frac{1}{2},m}}{h_{k+\frac{1}{2}}^x} - \frac{U_{z,k-\frac{1}{2},l+\frac{1}{2},m}}{h_{k-\frac{1}{2}}^x} \right) = \\
& \quad -V_{k,l+\frac{1}{2},m} \mathbf{F}_{y,k,l+\frac{1}{2},m}; \\
& \mathcal{S}_{k,l,m+\frac{1}{2}} \mathbf{E}_{z,k,l,m+\frac{1}{2}} - \left(\frac{U_{y,k+\frac{1}{2},l,m+\frac{1}{2}}}{h_{k+\frac{1}{2}}^x} - \frac{U_{y,k-\frac{1}{2},l,m+\frac{1}{2}}}{h_{k-\frac{1}{2}}^x} \right) + \left(\frac{U_{x,k,l+\frac{1}{2},m+\frac{1}{2}}}{h_{l+\frac{1}{2}}^y} - \frac{U_{x,k,l-\frac{1}{2},m+\frac{1}{2}}}{h_{l-\frac{1}{2}}^y} \right) = \\
& \quad -V_{k,l,m+\frac{1}{2}} \mathbf{F}_{z,k,l,m+\frac{1}{2}};
\end{aligned} \tag{6}$$

and

$$\begin{aligned}
\mathbf{e}_x(\mathbf{x}_r) &= \sum_{k,l,m} \mathcal{S}_{x,k+\frac{1}{2},l,m}^{ee}(\mathbf{x}_r) \mathbf{E}_{x,k+\frac{1}{2},l,m}; \\
\mathbf{e}_y(\mathbf{x}_r) &= \sum_{k,l,m} \mathcal{S}_{y,k,l+\frac{1}{2},m}^{ee}(\mathbf{x}_r) \mathbf{E}_{y,k,l+\frac{1}{2},m}; \\
\mathbf{e}_z(\mathbf{x}_r) &= \sum_{k,l,m} \mathcal{S}_{z,k,l,m+\frac{1}{2}}^{ee}(\mathbf{x}_r) \mathbf{E}_{z,k,l,m+\frac{1}{2}}; \\
\mathbf{h}_x(\mathbf{x}_r) &= \sum_{k,l,m} \mathcal{S}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^{hu}(\mathbf{x}_r) U_{x,k,l+\frac{1}{2},m+\frac{1}{2}}; \\
\mathbf{h}_y(\mathbf{x}_r) &= \sum_{k,l,m} \mathcal{S}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^{hu}(\mathbf{x}_r) U_{y,k+\frac{1}{2},l,m+\frac{1}{2}}; \\
\mathbf{h}_z(\mathbf{x}_r) &= \sum_{k,l,m} \mathcal{S}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^{hu}(\mathbf{x}_r) U_{z,k+\frac{1}{2},l+\frac{1}{2},m}.
\end{aligned} \tag{7}$$

The discrete values of the electric fields are interpreted as average quantities on the edges of the primal grid and the magnetic fields as average quantities on the faces of the primal grid. For instance, $\mathbf{E}_{x,k+\frac{1}{2},l,m}$ is the average of the x -component of the electric field on the edge $[x_k, x_{k+1}]$ and $U_{x,k,l+\frac{1}{2},m+\frac{1}{2}}$ the average on the face $[y_l, y_{l+1}] \times [z_m, z_{m+1}]$. \mathcal{S}^{ee} and \mathcal{S}^{hu} are sampling operators that map the fields to the receiver position \mathbf{x}_r . We have omitted references to the source, \mathbf{x}_s , to simplify the notation. In the first three equations, one needs μ_r^{-1} on the edges of the dual grid and in the three last equations, $\tilde{\sigma}$ on the edges of the primal grids. These values are evaluated by volume averaging. We define the cell volume by $V_{\alpha,\beta,\gamma} = h_\alpha^x h_\beta^y h_\gamma^z$ and let

$$\begin{aligned}
\mathcal{S}_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} &= \omega \mu_0 V_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} (\sigma_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} - \omega \varepsilon_{r,k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}}); \\
\mathcal{M}_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} &= V_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} \mu_{r,k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}}^{-1}.
\end{aligned}$$

The averaging to edges is accomplished by, for example,

$$\begin{aligned}
\mathcal{S}_{k+\frac{1}{2},l,m} &= \left(\mathcal{S}_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} + \mathcal{S}_{k+\frac{1}{2},l+\frac{1}{2},m-\frac{1}{2}} + \mathcal{S}_{k+\frac{1}{2},l-\frac{1}{2},m+\frac{1}{2}} + \mathcal{S}_{k+\frac{1}{2},l-\frac{1}{2},m-\frac{1}{2}} \right) / 4; \\
\mathcal{M}_{k,l+\frac{1}{2},m+\frac{1}{2}} &= \left(\mathcal{M}_{k+\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} + \mathcal{M}_{k-\frac{1}{2},l+\frac{1}{2},m+\frac{1}{2}} \right) / 2.
\end{aligned}$$

This implies that

$$\begin{aligned}\tilde{\sigma}_{k+\frac{1}{2},l,m} &= \frac{S_{k+\frac{1}{2},l,m}}{V_{k+\frac{1}{2},l,m}}; \\ \mu_{r,k,l+\frac{1}{2},m+\frac{1}{2}}^{-1} &= \frac{\mathcal{M}_{k,l+\frac{1}{2},m+\frac{1}{2}}}{V_{k,l+\frac{1}{2},m+\frac{1}{2}}}.\end{aligned}$$

We determine the discrete adjoint-state equations by equating to zero the derivatives of the augmented functional with respect the states variables [35, 36]. We do not list the augmented functional here, although the 1D case is described in Appendix B. The adjoint-state equations are ($\bar{\cdot}$ denotes the conjugate)

$$\begin{aligned}\frac{\bar{\lambda}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^U}{\mathcal{M}_{k,l+\frac{1}{2},m+\frac{1}{2}}} - \frac{\bar{\lambda}_{z,k,l+1,m+\frac{1}{2}}^{E,fit} - \bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fit}}{h_{l+\frac{1}{2}}^y} + \frac{\bar{\lambda}_{y,k,l+\frac{1}{2},m+1}^{E,fit} - \bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fit}}{h_{m+\frac{1}{2}}^z} &= \bar{\rho}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^U; \\ \frac{\bar{\lambda}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^U}{\mathcal{M}_{k+\frac{1}{2},l,m+\frac{1}{2}}} - \frac{\bar{\lambda}_{x,k+\frac{1}{2},l,m+1}^{E,fit} - \bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fit}}{h_{m+\frac{1}{2}}^z} + \frac{\bar{\lambda}_{z,k+1,l,m+\frac{1}{2}}^{E,fit} - \bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fit}}{h_{k+\frac{1}{2}}^x} &= \bar{\rho}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^U; \\ \frac{\bar{\lambda}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^U}{\mathcal{M}_{k+\frac{1}{2},l+\frac{1}{2},m}} - \frac{\bar{\lambda}_{y,k+1,l+\frac{1}{2},m}^{E,fit} - \bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fit}}{h_{k+\frac{1}{2}}^x} + \frac{\bar{\lambda}_{x,k+\frac{1}{2},l+1,m}^{E,fit} - \bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fit}}{h_{l+\frac{1}{2}}^y} &= \bar{\rho}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^U; \\ \mathcal{S}_{k+\frac{1}{2},l,m} \bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fit} - \left(\frac{\bar{\lambda}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^U}{h_{l+\frac{1}{2}}^y} - \frac{\bar{\lambda}_{z,k+\frac{1}{2},l-\frac{1}{2},m}^U}{h_{l-\frac{1}{2}}^y} \right) + \left(\frac{\bar{\lambda}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^U}{h_{m+\frac{1}{2}}^z} - \frac{\bar{\lambda}_{y,k+\frac{1}{2},l,m-\frac{1}{2}}^U}{h_{m-\frac{1}{2}}^z} \right) &= \bar{\rho}_{x,k+\frac{1}{2},l,m}^E; \\ \mathcal{S}_{k,l+\frac{1}{2},m} \bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fit} - \left(\frac{\bar{\lambda}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^U}{h_{m+\frac{1}{2}}^z} - \frac{\bar{\lambda}_{x,k,l+\frac{1}{2},m-\frac{1}{2}}^U}{h_{m-\frac{1}{2}}^z} \right) + \left(\frac{\bar{\lambda}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^U}{h_{k+\frac{1}{2}}^x} - \frac{\bar{\lambda}_{z,k-\frac{1}{2},l+\frac{1}{2},m}^U}{h_{k-\frac{1}{2}}^x} \right) &= \bar{\rho}_{y,k,l+\frac{1}{2},m}^E; \\ \mathcal{S}_{k,l,m+\frac{1}{2}} \bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fit} - \left(\frac{\bar{\lambda}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^U}{h_{k+\frac{1}{2}}^x} - \frac{\bar{\lambda}_{y,k-\frac{1}{2},l,m+\frac{1}{2}}^U}{h_{k-\frac{1}{2}}^x} \right) + \left(\frac{\bar{\lambda}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^U}{h_{l+\frac{1}{2}}^y} - \frac{\bar{\lambda}_{x,k,l-\frac{1}{2},m+\frac{1}{2}}^U}{h_{l-\frac{1}{2}}^y} \right) &= \bar{\rho}_{z,k,l,m+\frac{1}{2}}^E;\end{aligned}\tag{8}$$

with

$$\begin{aligned}\rho_{x,k+\frac{1}{2},l,m}^E(\mathbf{x}_r) &= \sum_r \mathbf{S}_{x,k+\frac{1}{2},l,m}^{ee} \frac{\partial F_c}{\partial e_x(\mathbf{x}_r)}; \\ \rho_{y,k,l+\frac{1}{2},m}^E(\mathbf{x}_r) &= \sum_r \mathbf{S}_{y,k,l+\frac{1}{2},m}^{ee} \frac{\partial F_c}{\partial e_y(\mathbf{x}_r)}; \\ \rho_{z,k,l,m+\frac{1}{2}}^E(\mathbf{x}_r) &= \sum_r \mathbf{S}_{z,k,l,m+\frac{1}{2}}^{ee} \frac{\partial F_c}{\partial e_z(\mathbf{x}_r)}; \\ \rho_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^U(\mathbf{x}_r) &= \sum_r \mathbf{S}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^{hu} \frac{\partial F_c}{\partial h_x(\mathbf{x}_r)}; \\ \rho_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^U(\mathbf{x}_r) &= \sum_r \mathbf{S}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^{hu} \frac{\partial F_c}{\partial h_y(\mathbf{x}_r)}; \\ \rho_{z,k+\frac{1}{2},l+\frac{1}{2},m}^U(\mathbf{x}_r) &= \sum_r \mathbf{S}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^{hu} \frac{\partial F_c}{\partial h_z(\mathbf{x}_r)}.\end{aligned}\tag{9}$$

The conjugate of the FIT adjoint states satisfies the FIT state system with specific source terms. The symmetry of the continuous operator is preserved (cf. [37]). For the implementation, this means that the same routines can be used for solving the state and the adjoint-state system, which is an advantage of the FIT over the FDM, see

Appendix A.

The discrete gradient for a single source at a single frequency is given by

$$\begin{aligned}
& - \sum_{k,l,m} \bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fit} \frac{\partial \mathcal{S}_{k+\frac{1}{2},l,m}}{\partial \mathbf{p}} \mathbf{E}_{x,k+\frac{1}{2},l,m} - \\
& \sum_{k,l,m} \bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fit} \frac{\partial \mathcal{S}_{k,l+\frac{1}{2},m}}{\partial \mathbf{p}} \mathbf{E}_{y,k,l+\frac{1}{2},m} - \\
& \sum_{k,l,m} \bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fit} \frac{\partial \mathcal{S}_{k,l,m+\frac{1}{2}}}{\partial \mathbf{p}} \mathbf{E}_{z,k,l,m+\frac{1}{2}}.
\end{aligned} \tag{10}$$

We have implemented the FIT discretization. This leads to a large linear system. We solved it with a preconditioned Krylov iterative method, BI-CGSTAB2, [19, 20]. The preconditioner for the linear iterative solver is a single multigrid cycle [22].

3. Model parameter scaling

The non-linear optimization of the misfit function is generally carried out with a non-linear conjugate gradient method, a Newton method, or a quasi-Newton method [38]. While the non-linear conjugate gradient or the Newton method is often used in geophysics [2, 14, 25, 39], we here decided to apply a quasi-Newton method. This approach is more suitable than the Newton method due to the size of the problem. It does not require the computation of the large sensitivity matrix (Fréchet derivatives). We adopted the BFGS method. This approach directly estimates the inverse of the Hessian matrix from the gradient vector of the misfit at each iteration. The size of the problem prevents us from working with the full Hessian. Therefore, we used the limited memory approach, where the inverse of the Hessian is approximated by q vectors of size M (M is the number of unknowns, i.e., the dimension of the model space) [40, 32, 33]. Hence, q controls the memory requirements of the approximation. We used the limited memory algorithm described in [34].

Although the method approximates the inverse of the Hessian, the convergence can still be improved with preconditioning. We chose to implement the preconditioner as a scaling of the unknowns, \mathbf{p} . With \mathbf{D} a regular matrix of $\mathbb{R}^{M \times M}$, we defined the scaled unknowns, $\tilde{\mathbf{p}}$, by

$$\tilde{\mathbf{p}} = \mathbf{D}^{-1} \mathbf{p}. \tag{11}$$

The misfit function becomes

$$\tilde{J}(\tilde{\mathbf{p}}) = J(\mathbf{D}\tilde{\mathbf{p}}) = J(\mathbf{p}), \tag{12}$$

the gradient

$$\nabla_{\tilde{\mathbf{p}}} \tilde{J}(\tilde{\mathbf{p}}) = \mathbf{D} \nabla_{\mathbf{p}} J(\mathbf{p}), \tag{13}$$

and the Hessian

$$\tilde{H}^J = \mathbf{D} \mathbf{H}^J \mathbf{D}, \tag{14}$$

where \mathbf{H}^J is the Hessian of J with respect to \mathbf{p} and $\tilde{\mathbf{H}}^J$ the Hessian of \tilde{J} with respect to $\tilde{\mathbf{p}}$.

The quasi-Newton update at iteration k is

$$\tilde{\mathbf{p}}^k = \tilde{\mathbf{p}}^{k-1} - \alpha^{k-1} \tilde{\mathbf{K}}^{k-1} \nabla_{\tilde{\mathbf{p}}} \tilde{J}(\tilde{\mathbf{p}}^{k-1}). \quad (15)$$

For the unknowns \mathbf{p} , this means

$$\mathbf{p}^k = \mathbf{p}^{k-1} - \alpha^{k-1} \mathbf{D} \tilde{\mathbf{K}}^{k-1} \mathbf{D} \nabla_{\mathbf{p}} J(\mathbf{p}^{k-1}). \quad (16)$$

Here, $\tilde{\mathbf{K}}^{k-1}$ is the BFGS approximation of the inverse of the Hessian $\tilde{\mathbf{H}}^J$ at iteration k . A good preconditioning matrix \mathbf{D} should be simple to compute and reduce the number of inversion iterations. For a quadratic misfit functional, this would be achieved by $\mathbf{D} = (\mathbf{H}^J)^{-1/2}$. For a non-quadratic, the Hessian depends on \mathbf{p} while we want a preconditioner matrix independent of \mathbf{p} .

The use of a preconditioner for the BFGS quasi-Newton method for electromagnetic data inversion was earlier studied by Newman and Boggs [30]. They proposed the use of the Gauss-Newton approximation of the Hessian, which requires the sensitivity matrix. Since it is too expensive to compute the sensitivity matrix at each iteration, they relied on a Born approximation and an estimate of the Green function in a simple reference model. For the low frequency electromagnetic applications in geophysics, the Born approximation is generally not really accurate and of limited use [5].

We then decided to test if a simple scaling of the unknowns would improve the convergence rate of the BFGS quasi-Newton inversion. We chose to work with a diagonal matrix \mathbf{D} for simplicity:

$$\mathbf{D}(x, y, z) = \begin{cases} 1 & \text{for } z \leq z_b; \\ [z^{-\gamma_m} \exp(-\beta_m \frac{z-z_b}{\delta}) + \varepsilon_r]^{-1} & \text{for } z > z_b; \end{cases} \quad (17)$$

with z_b a reference depth which may depend on x and y . For offshore CSEM inversion, z_b is the seafloor depth. Here, γ_m and β_m are two positive numbers and δ is a reference distance. We added the coefficient ε_r to account for the regularization term of equation 1. In practice, \mathbf{D} is also tapered for z larger than a given depth corresponding to about 2 to 3 skin depths to avoid boosting the noise. We called this scaling model depth weighting.

We proposed this expression for \mathbf{D} to crudely compensate for the amplitude exponential decay of the electromagnetic waves with the fact $\exp(-\beta_m \frac{z-z_b}{\delta})$ and for the geometrical spreading with the factor $z^{-\gamma_m}$. This means that δ can be the skin depth in an average resistivity and β_m around 1. We did not include a lateral dependency in \mathbf{D} because the least-squares functional contains a summation over the sources and receivers which are generally spread laterally.

4. Numerical results

4.1. Synthetic example

We applied the imaging approach with this depth weighting to a synthetic example based on a real geological setting [3]. The acquisition geometry is a line of 30 stations laid down on the sea floor. A horizontal electric source is towed at a distance of 40 m above the sea floor. A 2D inline section of the true resistivity model is displayed in figure 1.a. A crossline section is plotted in figure 5.a to show the 3D variations of the resistivity model. This model contains three main features: shallow resistors with resistivity between 10 to 20 Ohm.m, a deep 50 Ohm.m stratified reservoir, and a background resistivity increasing from 1.8 to 2.5 Ohm.m. For the data generation, we used a grid with a 10 m spacing in the z -direction around the sea floor as well as for the resistivity zones. Outside those zones, a 25 m spacing was used. We considered three frequencies: 0.25, 0.75, and 2.25 Hz. The stations were located between $x \simeq 10$ km and $x \simeq 40$ km. We used reciprocity to generate and process the data. This means that the data set corresponds to 90 source/frequency panels. This data set represents our “observed” data set.

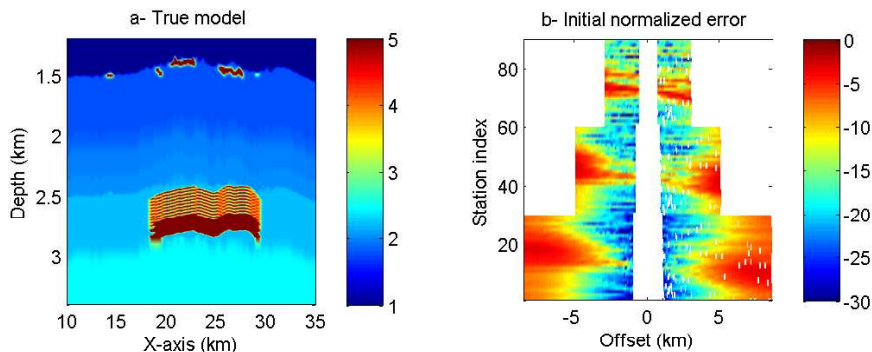


Figure 1. On the left, true resistivity model plotted in a linear scale. The resistivity values are clipped at 5 Ohm.m. On the right, the normalized misfit between the “observed” data, computed on the true model, and the synthetics computed on a 1.8 Ohm.m homogeneous earth. The scale is in dB.

For the imaging, the inversion grid has a 25 m spacing in depth. Therefore, the true resistivity is not in the model space of the inversion. For instance, the bathymetry depth difference between the inversion model and the true model can be as large as 15 m. During the runs, each frequency response was computed on a different computational grid to keep the number of points per skin depth constant [23]. We first computed the gradient on the computational grid and then used the adjoint of the projection operator from the inversion grid to the computational grid. In this way, we obtained the gradient on the inversion grid. This regridding approach introduces an error of few percent [23]. We chose not to use the same discretization for the data generation and for the inversion to make this synthetic study more meaningful by avoiding some of the issues of the “inverse crime”. We did not add noise to the data. However, the numerical noise

introduced by the regridding during the model generation and the inversion is different because the computational grids were different.

The initial model for the inversion has a constant resistivity of 1.8 Ohm.m below the sea floor and a water conductivity of 0.3 Ohm.m. We generated a synthetic set for this initial conductivity and computed the normalized errors between the synthetic and “observed” data in dB: $20 \log_{10} \left| \frac{e - e^{obs}}{e^{obs}} \right|$. The result is displayed in figure 1.b. The station indices from 1 to 30 correspond to a frequency of 0.25 Hz, between 31 and 60 to 0.75 Hz, and between 61 and 90 to 2.25 Hz. Data were absent in the white zones. Because of the diffusive nature of the electromagnetic waves, the low frequency signal diffuse further away than the high frequencies. The near-offset data, where the distance between the source and the receiver is small, are not taken into account in the inversion to avoid source representation effects. This is necessary when processing real data. At 0.25 Hz, we notice that the largest normalized errors occur around station indices 10 and 20, corresponding to x between about 20 and 30 km and offsets from 5 to 8 km. This suggests the presence of a deep resistive zone. The large normalized errors at high frequencies indicate the presence of shallow resistive zones.

The high frequency and short offset data are almost insensitive to the deep part of the model. Because we are often interested in the deep part and to compensate for the large amplitude decay versus offsets, the following data weighting is commonly applied:

$$W(\mathbf{x}_s, \mathbf{x}_r, \omega) = \frac{\|\mathbf{x}_s - \mathbf{x}_r\|^{\gamma_d}}{\omega^{\beta_f} \|E_1^0(\mathbf{x}_s, \mathbf{x}_r, \omega)\|^{\beta_d}}, \quad (18)$$

with positive constants γ_d , β_d , and β_f .

Smooth regularizations are not adequate for CSEM inversion because the goal is to retrieve thin resistors with high resistivity contrasts. We hence implemented the minimum norm support regularization [8]:

$$R(p) = \alpha_{nms} \sum_{klm} \frac{|p_{klm} - p_{klm}^{ref}|^2}{\beta_{nms}^2 + |p_{klm} - p_{klm}^{ref}|^2}, \quad (19)$$

where \mathbf{p}^{ref} is a reference model and α_{nms} and β_{nms} two positive numbers. The minimum support regularization should smooth \mathbf{p} for $|p_{klm} - p_{klm}^{ref}|$ smaller than about 2 to 3 β_{nms} ; however the large contrasts should not be penalized since the gradient of R should be small. In the inversion \mathbf{p} corresponds to the logarithm of resistivity and we chose $\beta_{nms} = 0.2$. The reference model is equal to the initial guess.

We took into account only the horizontal (in-line) electric responses to horizontal (in-line) emitting dipoles. We carried out four inversions with different data weights, W , and depth weighting, \mathbf{D} :

- D1: $\gamma_d = 0$ and $\beta_m = 0$,
- D2: $\gamma_d = 0.5$ and $\beta_m = 0$,
- D3: $\gamma_d = 0$, $\beta_m = 1$,
- D4: $\gamma_d = 0.5$, $\beta_m = 1$.

For all the inversions we have $\beta_d = 1$, $\beta_\omega = 0$ and $\gamma_m = 0$. In the examples, δ is the average skin depth at 0.25 Hz.

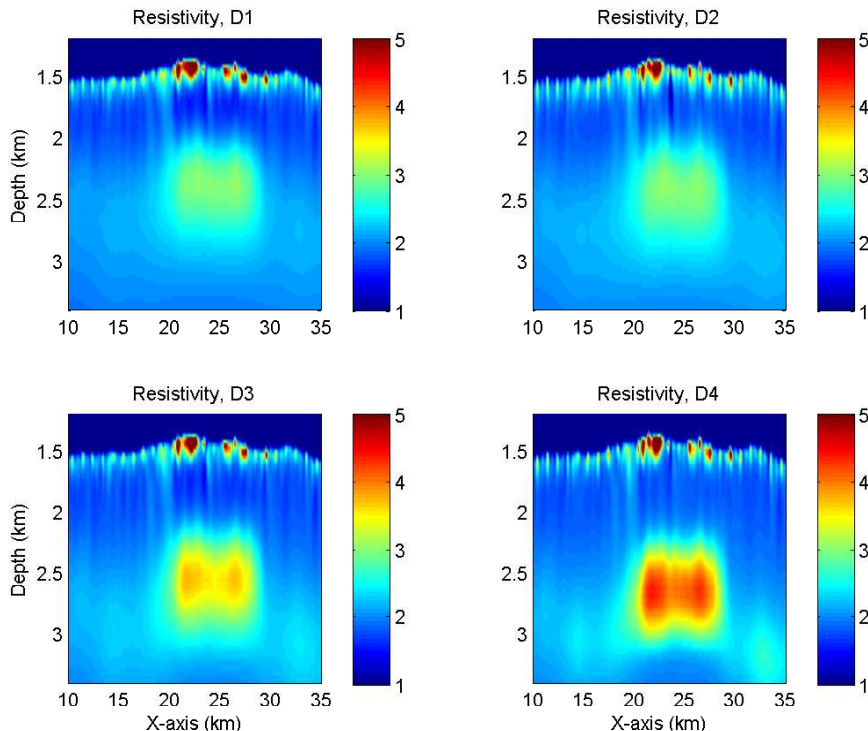


Figure 2. Optimal resistivity models at iteration 150 plotted in a linear scale. The resistivities are clipped at 5 Ohm.m.

The optimal resistivity models after 150 iterations for each depth and data weighting scheme are displayed in figure 2 with the weighted normalized errors in figure 3. The four inversion results well explain the data after 150 iterations. The normalized errors are around -20 dB, which is close to the noise floor in real applications. The four models mainly differ by the depth and the resistivity value of the deep resistive zone. This was expected because the depth resolution of the CSEM inversion is limited due to the diffusive effect and because of the non-uniqueness of the CSEM inversion. The deep resistive zone is however better retrieved with D4 and it is less well recovered with D1. We display in figure 5 a crossline section perpendicular to the inline section at $x = 22.4$ km. Due to the inline acquisition geometry, we cannot retrieve the offline shallow resistors and we notice the so-called “inversion smiles”. We display the decrease of the weighted least-squares misfit versus the number of iterations in figure 4.a. We obtained the fastest decrease for D4. The three other depth and data weighting gave a relatively similar misfit decrease. We carried out those inversions with $q = 1$, which limits the memory allocated for the approximated BFGS inverse of the Hessian to one vector. We also inverted the CSEM data with the depth and data weighting D4 with $q = 3$ and $q = 5$. The convergence histories are plotted in figure 4: the convergence did not improve when increasing the number of vectors that determines the approximation of the BFGS inverse of the Hessian.

To further understand the relevance of the depth and data weighting, we display

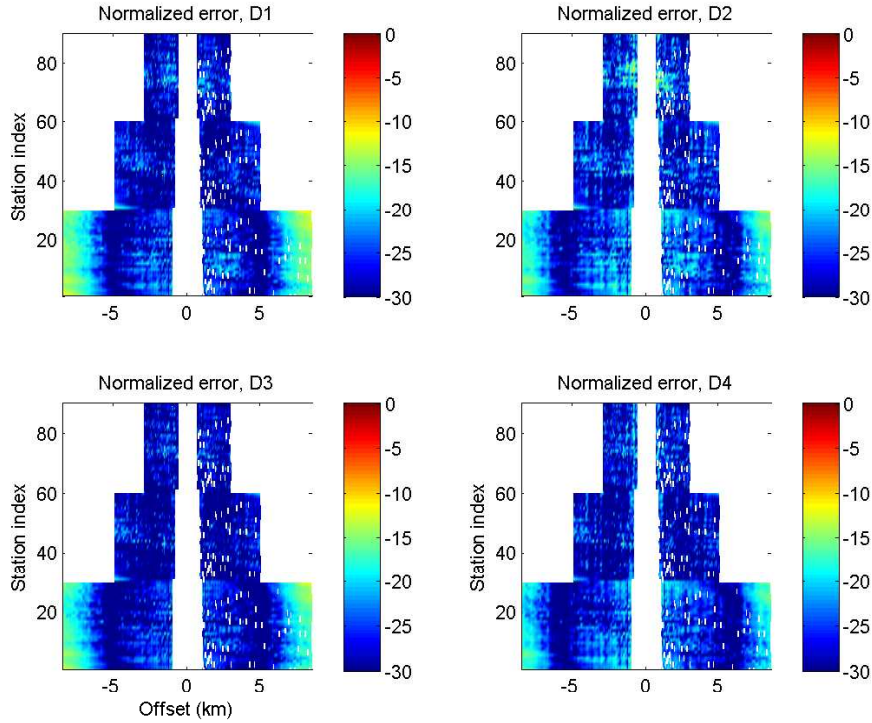


Figure 3. Weighted normalized errors for the four depth-weighting schemes plotted in a dB scale (the weighted normalized error is $\|W(\mathbf{x}_s, \mathbf{x}_r, f)(e_{x,s,r,f}(\mathbf{p}) - e_{x,s,r,f}^{obs})\|$). The white spots correspond to missing data. We removed some of the data samples to mimic a real situation where some data samples are too noisy to be included in the inversion due to instrument noise.

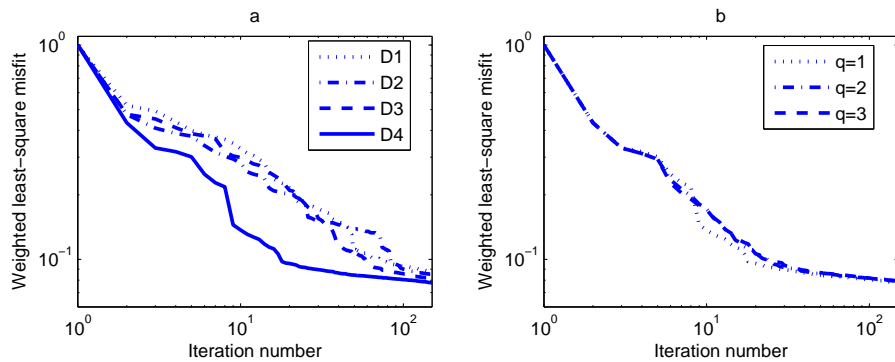


Figure 4. On the left, the decrease of the weighted least-squares misfit versus the number of iterations for the four depth and data weighting schemes when the approximated BFGS inverse of the Hessian has a size of one vector ($q = 1$). On the right, the decrease of the weighted least-square with the depth and data weighting D4 for three different sizes of the approximated BFGS inverse of the Hessian.

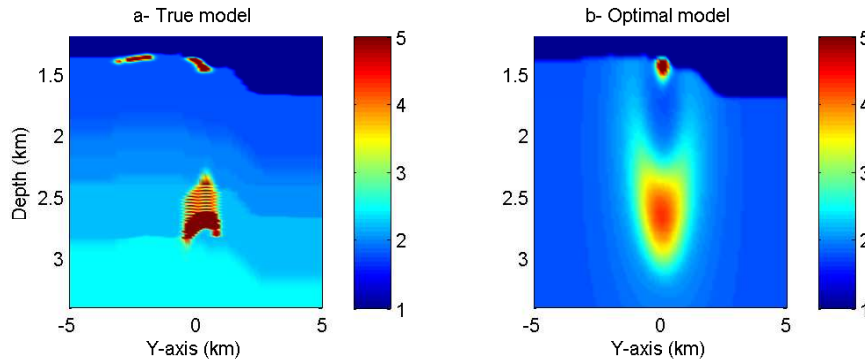


Figure 5. Crossline sections at $x = 22.5$ km of the true resistivity model (a) and of the optimal resistivity model after 150 iterations of the inversion D4 (b). The resistivities are clipped at 5 Ohm.m.

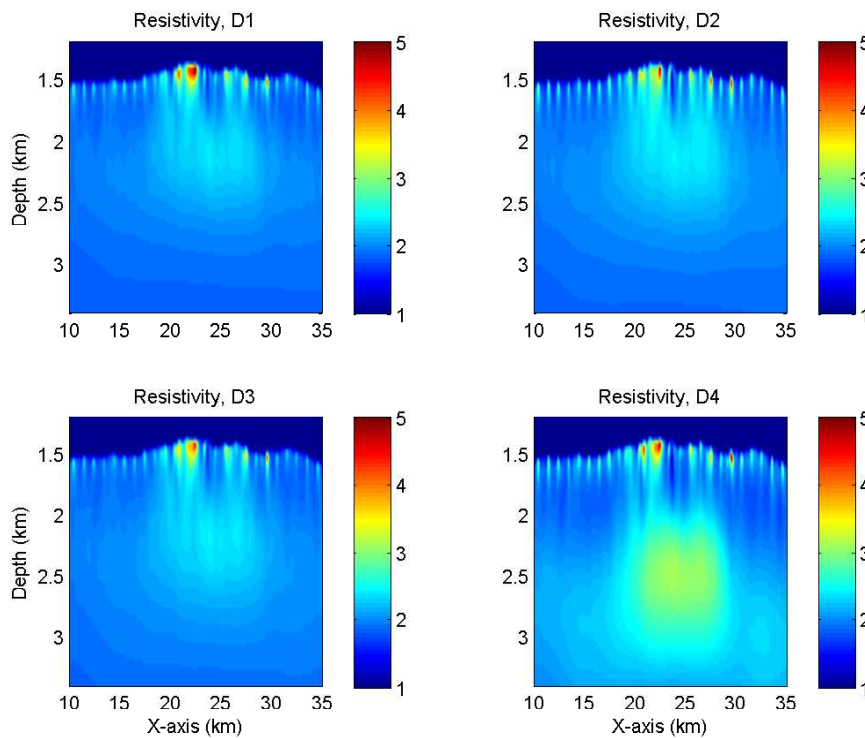


Figure 6. Optimal resistivity models at iteration 20. The resistivities are clipped at 5 Ohm.m.

the results of the four inversions after 20 iterations in figure 6 and after 50 iterations in figure 7. We can clearly see that with the depth and data weighting D4, the deeper part is updated more rapidly. This speeds up the convergence significantly. Although the depth weighting is crude, it helps to balance the shallow and deep resistivity updates during the iterations. The inversions were computed on 30 cores (AMD Operton 2.4 Ghz). One iteration took a bit more than 18 minutes. It requires the computation of 90 forward and 90 backward modeling steps. This means that one evaluation of system 6 took a

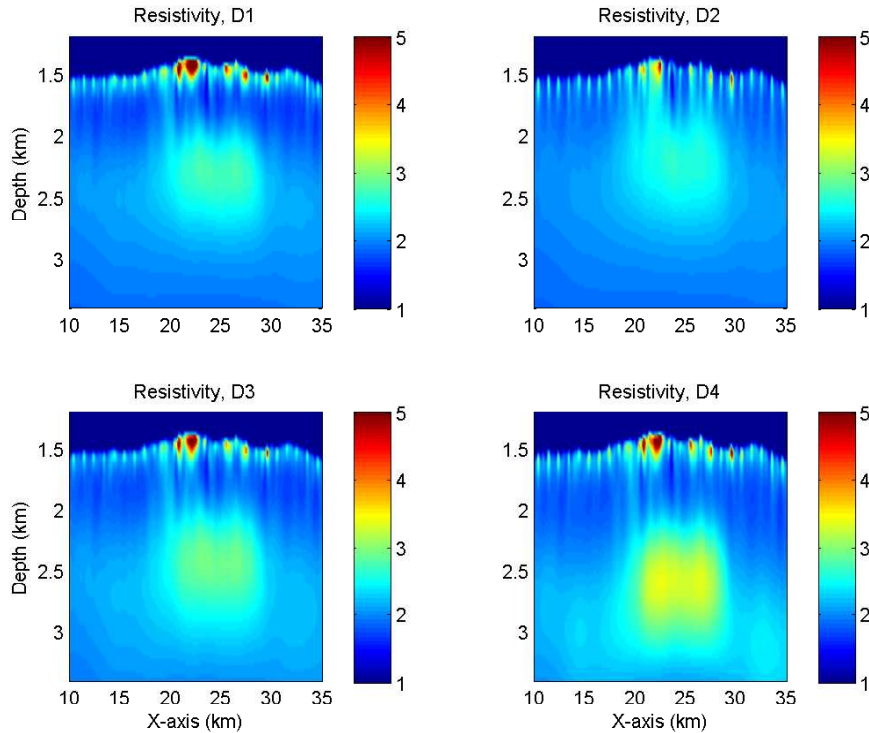


Figure 7. Optimal resistivity models at iteration 50. The resistivities are clipped at 5 Ohm.m.

bit more than 3 minutes on a single core, which shows the efficiency of the modeling approach for deep-water CSEM applications. For each computation, the grid contained about 117 by 69 by 130 points (a total of about 1 million points) while the inversion grid contained about 3.5 million points. If we consider the fact that the inversion D4 could have been stopped at iteration 50, the inversion computation time was about 16 hours. We observe in figure 4.a that after iteration 20, the convergence rate significantly decreases for the case D4. This is probably due to the regularization, since $R(\mathbf{p})$ stays around 0.04 after iteration 10.

Those results show that the combination of depth weighting (preconditioner) with a data weighting that emphasizes the large offsets gave the best result. Boosting the large offsets in CSEM inversion is common practice because it was noticed that the information on the deeper part of the earth is mainly contained in the large offset data. In a similar way, we can also increase the role of the low frequencies in the misfit function. To illustrate the role of the frequency weighting in the misfit data, we repeated the four inversions but with $\beta_f = 0.25$. The resistivities at iteration 20, 50, and 100 are displayed in figures 8, 9, and 10, respectively. We notice that the target zone is updated earlier with a frequency data weighting than without. Once again, the combination of depth and data weighting gave the best result. The convergence history is plotted in figure 11.

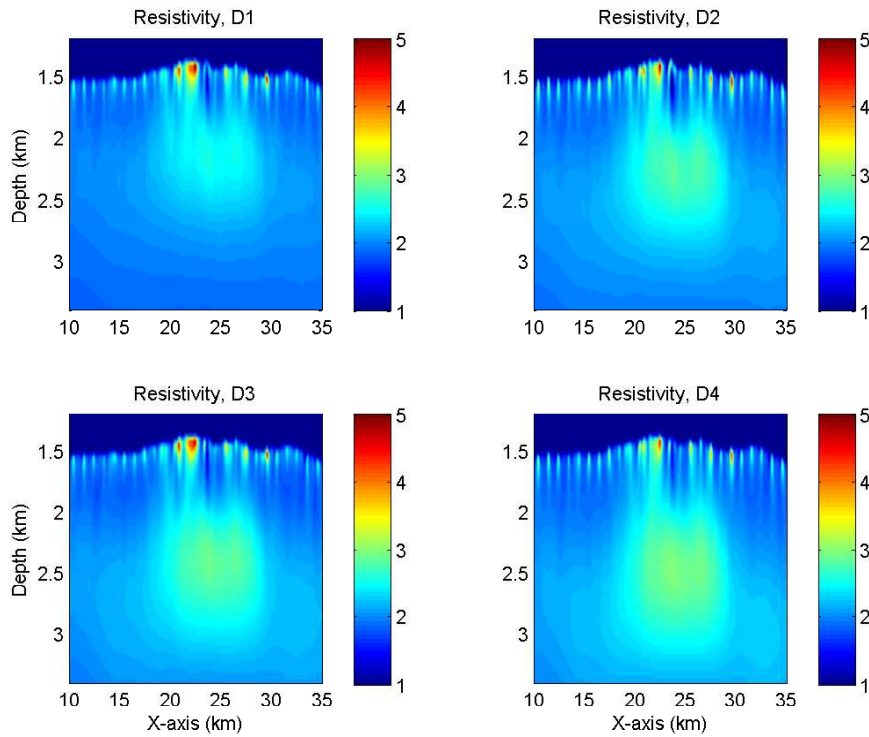


Figure 8. Optimal resistivity models at iteration 20 with frequency data weighting in the inversion. The resistivities are clipped at 5 Ohm.m.

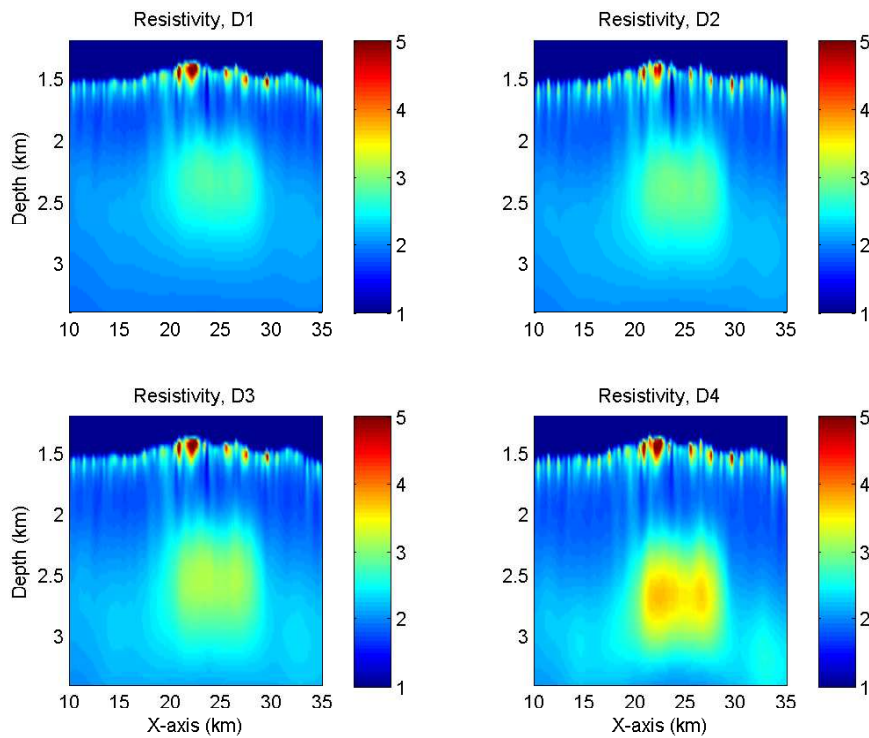


Figure 9. Optimal resistivity models at iteration 50 with frequency data weighting in the inversion. The resistivities are clipped at 5 Ohm.m.

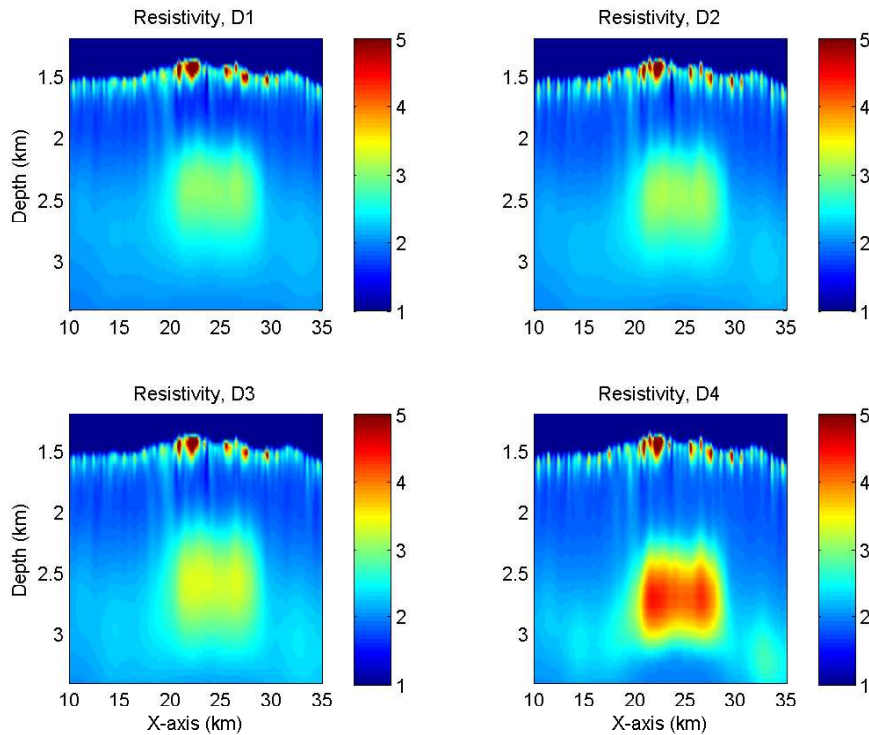


Figure 10. Optimal resistivity models at iteration 100 with frequency data weighting in the inversion. The resistivities are clipped at 5 Ohm.m.

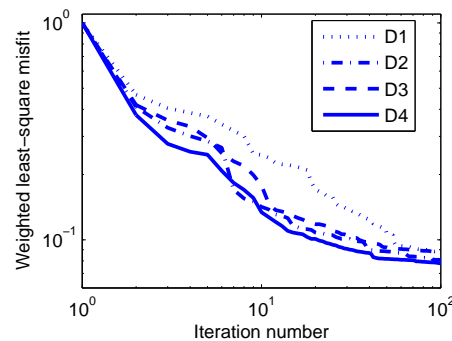


Figure 11. Convergence history of the inversions with frequency data weighting in the inversion.

4.2. Results with Troll data set

We now invert the Troll data set with the depth weighting D_4 . The data have been recorded over the North Sea Troll west gas field in 2003. Imaging results of this data set can be found, for instance, in [6]. The geology is fairly simple and the reservoir has a high resistivity. This is an ideal case for CSEM application. The water depth is about 300 m and the reservoir is about 1500 m deep. The acquisition geometry is a standard line acquisition with 24 receivers. We inverted the inline electric responses of the inline

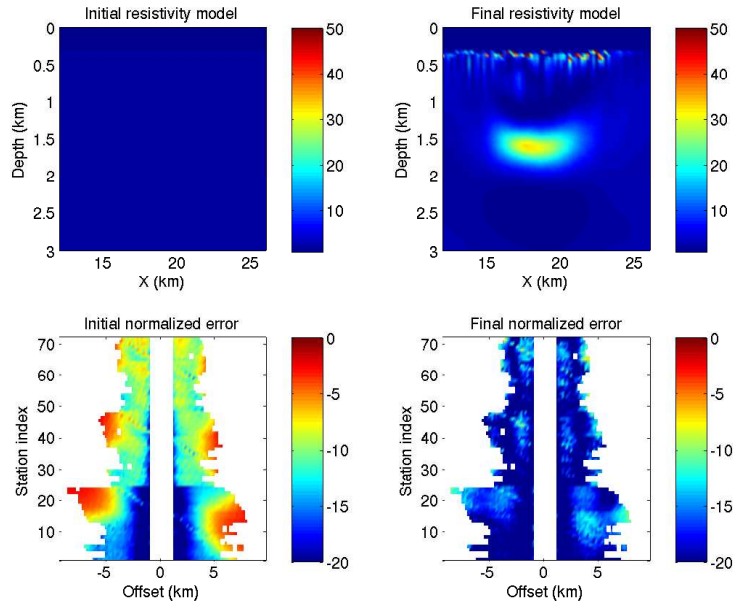


Figure 12. Inversion results of the Troll data set. The resistivity values (in Ohm-m) are clipped at 25 Ohm-m. The normalized misfit (in db) between the observed data and the synthetics. The station indices from 1 to 24 correspond to the data at 0.25 Hz, from 35 to 48 to the data at 0.75 Hz, and from 49 to 72 to the data at 1.25 Hz.

electric dipole at 0.25, 0.75 and 1.25 Hz. The water resistivity is fixed at 0.3 Ohm-m and the initial earth resistivity is 2.5 Ohm-m. In Figure 12, we plotted the initial and final resistivity models and the initial and final normalized data misfit. On the initial normalized data plot, we can see a large anomaly at 0.25 Hz (station index from 1 to 24). This anomaly is visible at offsets larger than 2 km and is hardly present at 1.25 Hz (station index from 49 to 72). This is an indication of a deep resistive layer. On the initial normalized data plot at 1.25 Hz, we can also notice that either the 2.5 Ohm-m initial resistivity does not fully explain the data indicating some shallow variations or some calibration issues of the data still exist due to noise. The final resistivity model clearly displays the Troll gas field at the correct depth. Some shallow variations are also present. The final normalized data plot shows that this final model correctly interprets the data. We plotted the result up to 3 km depth to illustrate that the use of the depth weighting did not create artefacts in the final resistivity model.

5. Conclusions

We discussed some computational aspects of the resistivity inversion from controlled-source electromagnetic data and we considered a realistic and rather complicated 3D numerical example and a simple real data example. The implementation is simpler when we discretize Maxwell's equation with the Finite-Integration Technique because the discrete operator retains the symmetry property of the continuous operator. We did not try to evaluate an approximation of the Hessian in a reference model because the

Born approximation is generally of limited use with the low frequency electromagnetic applications in geophysics. With the inverse of the Hessian estimated by the BFGS method, we numerically showed that a simple exponential depth weighting combined with proper data weighting helps to balance the shallow and deep part of the resistivity update and provides a significant improvement of the inversion convergence and inversion results. Although this approach is based on simple heuristics, it allows us to obtain relevant resistivity images from CSEM data in about 50 iterations, which is a quite satisfactory result. This reduces the number of iterations by 2 to 4 compared to an inversion without depth weighting. We did not compare our simple approach with an inversion based on a Gauss-Newton approximation of the Hessian.

Acknowledgments

The authors thank Statoil, EMGS, Norks Hydro, Petoro, Norske Shell, Total, and ConocoPhillips for the Troll data set. The authors also thanks Peter van der Smán, Shell International E&P, for the preprocessing of the frequency-domain Troll data set.

Appendix A. The adjoint-state equations for the FDM and the FIT

The FDM proposed by Yee [18] can be extended to stretched grid. The electric field is discretized at the midpoints of the edges on the primal grid. We denote its components by $\mathbf{E}_{x,k+\frac{1}{2},l,m}$, $\mathbf{E}_{y,k,l+\frac{1}{2},m}$, and $\mathbf{E}_{z,k,l,m+\frac{1}{2}}$. This notation means that, for instance, the x -component of \mathbf{E} is discretized at the point $(x_{k+\frac{1}{2}}, y_l, z_m)$. The magnetic field is discretized at the midpoints of the edges on the dual grid. These correspond to the centers of the faces of the primal grid. We denote those components by $\mathbf{H}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}$, $\mathbf{H}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}$, and $\mathbf{H}_{z,k+\frac{1}{2},l+\frac{1}{2},m}$. Using centered finite-differences, the discretization of equation (3), $\mathbf{W} = \mu_r^{-1} \nabla \times \mathbf{E}$, becomes

$$\begin{aligned}
& \frac{V_{k,l+\frac{1}{2},m+\frac{1}{2}}}{\mathcal{M}_{k,l+\frac{1}{2},m+\frac{1}{2}}} \mathbf{W}_{x,k,l+\frac{1}{2},m+\frac{1}{2}} - \left(\frac{\mathbf{E}_{z,k,l+1,m+\frac{1}{2}} - \mathbf{E}_{z,k,l,m+\frac{1}{2}}}{h_{l+\frac{1}{2}}^y} - \frac{\mathbf{E}_{y,k,l+\frac{1}{2},m+1} - \mathbf{E}_{y,k,l+\frac{1}{2},m}}{h_{m+\frac{1}{2}}^z} \right) = 0; \\
& \frac{V_{k+\frac{1}{2},l,m+\frac{1}{2}}}{\mathcal{M}_{k+\frac{1}{2},l,m+\frac{1}{2}}} \mathbf{W}_{y,k+\frac{1}{2},l,m+\frac{1}{2}} - \left(\frac{\mathbf{E}_{x,k+\frac{1}{2},l,m+1} - \mathbf{E}_{x,k+\frac{1}{2},l,m}}{h_{m+\frac{1}{2}}^z} - \frac{\mathbf{E}_{z,k+1,l,m+\frac{1}{2}} - \mathbf{E}_{z,k,l,m+\frac{1}{2}}}{h_{k+\frac{1}{2}}^x} \right) = 0; \\
& \frac{V_{k+\frac{1}{2},l+\frac{1}{2},m}}{\mathcal{M}_{k+\frac{1}{2},l+\frac{1}{2},m}} \mathbf{W}_{z,k+\frac{1}{2},l+\frac{1}{2},m} - \left(\frac{\mathbf{E}_{y,k+1,l+\frac{1}{2},m} - \mathbf{E}_{y,k,l+\frac{1}{2},m}}{h_{k+\frac{1}{2}}^x} - \frac{\mathbf{E}_{x,k+\frac{1}{2},l+1,m} - \mathbf{E}_{x,k+\frac{1}{2},l,m}}{h_{l+\frac{1}{2}}^y} \right) = 0; \\
& \frac{S_{k+\frac{1}{2},l,m}}{V_{k+\frac{1}{2},l,m}} \mathbf{E}_{x,k+\frac{1}{2},l,m} - \left(\frac{\mathbf{W}_{z,k+\frac{1}{2},l+\frac{1}{2},m} - \mathbf{W}_{z,k+\frac{1}{2},l-\frac{1}{2},m}}{h_l^y} - \frac{\mathbf{W}_{y,k+\frac{1}{2},l,m+\frac{1}{2}} - \mathbf{W}_{y,k+\frac{1}{2},l,m-\frac{1}{2}}}{h_m^z} \right) = \\
& \qquad \qquad \qquad - \mathbf{F}_{x,k+\frac{1}{2},l,m}; \\
& \frac{S_{k,l+\frac{1}{2},m}}{V_{k,l+\frac{1}{2},m}} \mathbf{E}_{y,k,l+\frac{1}{2},m} - \left(\frac{\mathbf{W}_{x,k,l+\frac{1}{2},m+\frac{1}{2}} - \mathbf{W}_{x,k,l+\frac{1}{2},m-\frac{1}{2}}}{h_m^z} - \frac{\mathbf{W}_{z,k+\frac{1}{2},l+\frac{1}{2},m} - \mathbf{W}_{z,k-\frac{1}{2},l+\frac{1}{2},m}}{h_k^x} \right) = \\
& \qquad \qquad \qquad - \mathbf{F}_{y,k,l+\frac{1}{2},m}; \\
& \frac{S_{k,l,m+\frac{1}{2}}}{V_{k,l,m+\frac{1}{2}}} \mathbf{E}_{z,k,l,m+\frac{1}{2}} - \left(\frac{\mathbf{W}_{y,k+\frac{1}{2},l,m+\frac{1}{2}} - \mathbf{W}_{y,k-\frac{1}{2},l,m+\frac{1}{2}}}{h_k^x} - \frac{\mathbf{W}_{x,k,l+\frac{1}{2},m+\frac{1}{2}} - \mathbf{W}_{x,k,l-\frac{1}{2},m+\frac{1}{2}}}{h_l^y} \right) = \\
& \qquad \qquad \qquad - \mathbf{F}_{z,k,l,m+\frac{1}{2}};
\end{aligned} \tag{A.1}$$

and

$$\begin{aligned}
\mathbf{e}_x(\mathbf{x}_r) &= \sum_{k,l,m} \mathbf{S}_{x,k+\frac{1}{2},l,m}^{ee}(\mathbf{x}_r) \mathbf{E}_{x,k+\frac{1}{2},l,m}; \\
\mathbf{e}_y(\mathbf{x}_r) &= \sum_{k,l,m} \mathbf{S}_{y,k,l+\frac{1}{2},m}^{ee}(\mathbf{x}_r) \mathbf{E}_{y,k,l+\frac{1}{2},m}; \\
\mathbf{e}_z(\mathbf{x}_r) &= \sum_{k,l,m} \mathbf{S}_{z,k,l,m+\frac{1}{2}}^{ee}(\mathbf{x}_r) \mathbf{E}_{z,k,l,m+\frac{1}{2}}; \\
\mathbf{h}_x(\mathbf{x}_r) &= \sum_{k,l,m} \mathbf{S}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^{hw}(\mathbf{x}_r) \mathbf{W}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}; \\
\mathbf{h}_y(\mathbf{x}_r) &= \sum_{k,l,m} \mathbf{S}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^{hw}(\mathbf{x}_r) \mathbf{W}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}; \\
\mathbf{h}_z(\mathbf{x}_r) &= \sum_{k,l,m} \mathbf{S}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^{hw}(\mathbf{x}_r) \mathbf{W}_{z,k+\frac{1}{2},l+\frac{1}{2},m}.
\end{aligned} \tag{A.2}$$

The FDM and FIT schemes (A.1) and (6) are equivalent. Both have second-order accuracy [37]. Also, the average quantities and the midpoint values are the same up to the second-order. Hence, by replacing $\mathbf{U}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}$ with $V_{k,l+\frac{1}{2},m+\frac{1}{2}} \mathbf{W}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}$, $\mathbf{U}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}$ with $V_{k+\frac{1}{2},l,m+\frac{1}{2}} \mathbf{W}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}$, and $\mathbf{U}_{z,k+\frac{1}{2},l+\frac{1}{2},m}$ with $V_{k+\frac{1}{2},l+\frac{1}{2},m} \mathbf{W}_{z,k+\frac{1}{2},l+\frac{1}{2},m}$ in system (6), we obtain the system (A.1). From a forward modeling point of view, there are *a priori* no reasons to prefer one scheme above the other. The situation is different for the imaging problem because we also need to solve the adjoint system. The continuous electromagnetic operator is symmetric, meaning that the conjugate of the adjoint-state variables satisfies the same equation as the state variables but for different source terms. Equating to zeros the derivatives of the augmented functional with respect to the state variables \mathbf{W} and \mathbf{E} for the FDM leads to the backward system

$$\left\{ \begin{aligned}
& \frac{V_{k,l+\frac{1}{2},m+\frac{1}{2}}}{\mathcal{M}_{k,l+\frac{1}{2},m+\frac{1}{2}}} \bar{\lambda}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^W - \left(\frac{\bar{\lambda}_{z,k,l+1,m+\frac{1}{2}}^{E,fd}}{h_{l+1}^y} - \frac{\bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fd}}{h_l^y} \right) - \left(\frac{\bar{\lambda}_{y,k,l+\frac{1}{2},m+1}^{E,fd}}{h_{m+1}^z} - \frac{\bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fd}}{h_m^z} \right) = \\
& \qquad \qquad \qquad \bar{\rho}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^W; \\
& \frac{V_{k+\frac{1}{2},l,m+\frac{1}{2}}}{\mathcal{M}_{k+\frac{1}{2},l,m+\frac{1}{2}}} \bar{\lambda}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^W - \left(\frac{\bar{\lambda}_{x,k+\frac{1}{2},l,m+1}^{E,fd}}{h_{m+1}^z} - \frac{\bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fd}}{h_m^z} \right) - \left(\frac{\bar{\lambda}_{z,k+1,l,m+\frac{1}{2}}^{E,fd}}{h_{k+1}^x} - \frac{\bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fd}}{h_k^x} \right) = \\
& \qquad \qquad \qquad \bar{\rho}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^W; \\
& \frac{V_{k+\frac{1}{2},l+\frac{1}{2},m}}{\mathcal{M}_{k+\frac{1}{2},l+\frac{1}{2},m}} \bar{\lambda}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^W - \left(\frac{\bar{\lambda}_{y,k+1,l+\frac{1}{2},m}^{E,fd}}{h_{k+1}^x} - \frac{\bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fd}}{h_k^x} \right) - \left(\frac{\bar{\lambda}_{x,k+\frac{1}{2},l+1,m}^{E,fd}}{h_{l+1}^y} - \frac{\bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fd}}{h_l^y} \right) = \\
& \qquad \qquad \qquad \bar{\rho}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^W; \\
& \frac{S_{k+\frac{1}{2},l,m}}{V_{k+\frac{1}{2},l,m}} \bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fd} - \left(\frac{\bar{\lambda}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^W}{h_{l+\frac{1}{2}}^y} - \frac{\bar{\lambda}_{z,k+\frac{1}{2},l-\frac{1}{2},m}^W}{h_{l-\frac{1}{2}}^y} \right) - \left(\frac{\bar{\lambda}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^W}{h_{m+\frac{1}{2}}^z} - \frac{\bar{\lambda}_{y,k+\frac{1}{2},l,m-\frac{1}{2}}^W}{h_{m-\frac{1}{2}}^z} \right) = \\
& \qquad \qquad \qquad \bar{\rho}_{x,k+\frac{1}{2},l,m}^E; \\
& \frac{S_{k,l+\frac{1}{2},m}}{V_{k,l+\frac{1}{2},m}} \bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fd} - \left(\frac{\bar{\lambda}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^W}{h_{m+\frac{1}{2}}^z} - \frac{\bar{\lambda}_{x,k,l+\frac{1}{2},m-\frac{1}{2}}^W}{h_{m-\frac{1}{2}}^z} \right) - \left(\frac{\bar{\lambda}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^W}{h_{k+\frac{1}{2}}^x} - \frac{\bar{\lambda}_{z,k-\frac{1}{2},l+\frac{1}{2},m}^W}{h_{k-\frac{1}{2}}^x} \right) = \\
& \qquad \qquad \qquad \bar{\rho}_{y,k,l+\frac{1}{2},m}^E; \\
& \frac{S_{k,l,m+\frac{1}{2}}}{V_{k,l,m+\frac{1}{2}}} \bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fd} - \left(\frac{\bar{\lambda}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^W}{h_{k+\frac{1}{2}}^x} - \frac{\bar{\lambda}_{y,k-\frac{1}{2},l,m+\frac{1}{2}}^W}{h_{k-\frac{1}{2}}^x} \right) - \left(\frac{\bar{\lambda}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^W}{h_{l+\frac{1}{2}}^y} - \frac{\bar{\lambda}_{x,k,l-\frac{1}{2},m+\frac{1}{2}}^W}{h_{l-\frac{1}{2}}^y} \right) = \\
& \qquad \qquad \qquad \bar{\rho}_{z,k,l,m+\frac{1}{2}}^E;
\end{aligned} \right. \tag{A.3}$$

with

$$\begin{aligned}
\rho_x^E(\mathbf{x}_r) &= \sum_r \mathbf{S}_{x,k+\frac{1}{2},l,m}^{ee} \frac{\partial F_c}{\partial e_x(\mathbf{x}_r)}; \\
\rho_y^E(\mathbf{x}_r) &= \sum_r \mathbf{S}_{y,k,l+\frac{1}{2},m}^{ee} \frac{\partial F_c}{\partial e_y(\mathbf{x}_r)}; \\
\rho_z^E(\mathbf{x}_r) &= \sum_r \mathbf{S}_{z,k,l,m+\frac{1}{2}}^{ee} \frac{\partial F_c}{\partial e_z(\mathbf{x}_r)}; \\
\rho_x^W(\mathbf{x}_r) &= \sum_r \mathbf{S}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^{hw} \frac{\partial F_c}{\partial h_x(\mathbf{x}_r)}; \\
\rho_y^W(\mathbf{x}_r) &= \sum_r \mathbf{S}_{y,k+\frac{1}{2},l,m+\frac{1}{2}}^{hw} \frac{\partial F_c}{\partial h_y(\mathbf{x}_r)}; \\
\rho_z^W(\mathbf{x}_r) &= \sum_r \mathbf{S}_{z,k+\frac{1}{2},l+\frac{1}{2},m}^{hw} \frac{\partial F_c}{\partial h_z(\mathbf{x}_r)}.
\end{aligned} \tag{A.4}$$

We observe that the discrete FDM adjoint-state system is different from the FDM state system for stretched grids. Clearly, the FDM operator is not symmetric.

The discrete gradient for a single source at a single frequency is given by

$$\begin{aligned}
& - \sum_{k,l,m} \left[\frac{\bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fd}}{V_{k+\frac{1}{2},l,m}} \frac{\partial \mathcal{S}_{k+\frac{1}{2},l,m}}{\partial \mathbf{p}} \mathbf{E}_{x,k+\frac{1}{2},l,m} - \frac{\bar{\lambda}_{y,k,l+\frac{1}{2},m}^{E,fd}}{V_{k,l+\frac{1}{2},m}} \frac{\partial \mathcal{S}_{k,l+\frac{1}{2},m}}{\partial \mathbf{p}} \mathbf{E}_{y,k,l+\frac{1}{2},m} - \right. \\
& \left. \frac{\bar{\lambda}_{z,k,l,m+\frac{1}{2}}^{E,fd}}{V_{k,l,m+\frac{1}{2}}} \frac{\partial \mathcal{S}_{k,l,m+\frac{1}{2}}}{\partial \mathbf{p}} \mathbf{E}_{z,k,l,m+\frac{1}{2}} \right].
\end{aligned} \tag{A.5}$$

The FDM and FIT gradient formulas (A.5) and (10) provide, of course, the same result since $\bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fit} = \frac{\bar{\lambda}_{x,k+\frac{1}{2},l,m}^{E,fd}}{V_{k+\frac{1}{2},l,m}}$, $\bar{\lambda}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^W = \bar{\lambda}_{x_k,l+\frac{1}{2},m+\frac{1}{2}}^U$, and $\bar{\rho}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^W = V_{k,l+\frac{1}{2},m+\frac{1}{2}} \bar{\rho}_{x,k,l+\frac{1}{2},m+\frac{1}{2}}^U$. However, having a symmetric discrete operator simplifies the implementation because the forward and background system can be implemented with the same routines.

Appendix B. The 1D case

In this appendix we present the complete gradient derivation with the finite-difference method and the discretization based on the Finite-Integration Technique for a 1D problem. Then, we evaluate the errors we would commit if we approximate the transpose of the adjoint-state system with the forward system in the FDM. We thereby illustrate the well-known fact that we should apply the adjoint-state method to the discrete equations and not discretize the adjoint-state equations found from the continuous case.

The 1D partial differential equation on $\Omega \subset \mathbb{R}$ reads

$$\frac{\partial^2 u}{\partial x^2} + \omega \mu_0 \sigma u = f, \tag{B.1}$$

with $u = 0$ on Γ the boundary points on Ω .

The FDM gives

$$\frac{u_{k+1} - u_k}{h_k h_{k+\frac{1}{2}}} - \frac{u_k - u_{k-1}}{h_k h_{k-\frac{1}{2}}} + \omega \mu_0 \sigma_k u_k = f_k. \tag{B.2}$$

The FIT discretization is

$$\frac{u_{k+1} - u_k}{h_{k+\frac{1}{2}}} - \frac{u_k - u_{k-1}}{h_{k-\frac{1}{2}}} + \omega \mu_0 h_k \sigma_k u_k = h_k f_k, \tag{B.3}$$

with $u_0 = u_n = 0$. We consider the least-squares functional

$$f(\sigma) = \frac{1}{2} \sum_{k=1}^{n-1} \|u_k - u_k^{obs}\|^2, \quad (\text{B.4})$$

where u_k^{obs} are data recorded at x_k .

The equation for the continuous adjoint states is

$$\frac{\partial^2 \bar{\lambda}}{\partial x^2} + i\omega\mu_0\sigma\bar{\lambda} = \overline{u - u^{obs}}. \quad (\text{B.5})$$

The Lagrangian associated with the FDM is

$$\begin{aligned} \mathcal{L}^a(\tilde{u}_k, \tilde{\lambda}_k) = \\ \frac{1}{2} \sum_{k=1}^{n-1} \|\tilde{u}_k - u_k^{obs}\|^2 - \sum_{k=1}^{n-1} \tilde{\lambda}_k \left(\frac{\tilde{u}_{k+1} - \tilde{u}_k}{h_k h_{k+\frac{1}{2}}} - \frac{\tilde{u}_k - \tilde{u}_{k-1}}{h_k h_{k-\frac{1}{2}}} + i\omega\mu_0\sigma_k u_k - f_k \right). \end{aligned} \quad (\text{B.6})$$

The derivative of \mathcal{L}^a with respect to \tilde{u}_k provides the adjoint-state equation of equation (B.2)

$$\frac{1}{h_{k+\frac{1}{2}}} \left(\frac{\bar{\lambda}_{k+1}^a}{h_{k+1}} - \frac{\bar{\lambda}_k^a}{h_k} \right) - \frac{1}{h_{k-\frac{1}{2}}} \left(\frac{\bar{\lambda}_k^a}{h_k} - \frac{\bar{\lambda}_{k-1}^a}{h_{k-1}} \right) + i\omega\mu_0\sigma_k \bar{\lambda}_k^a = \overline{u_k - u_k^d} \quad (\text{B.7})$$

with $\lambda_0 = \lambda_n = 0$.

Similarly the Lagrangian associated with the FIT discretization is

$$\begin{aligned} \mathcal{L}^a(\tilde{u}_k, \tilde{\lambda}_k) = \\ \frac{1}{2} \sum_{k=1}^{n-1} \|\tilde{u}_k - u_k^{obs}\|^2 - \sum_{k=1}^{n-1} \tilde{\lambda}_k \left(\frac{u_{k+1} - u_k}{h_{k+\frac{1}{2}}} - \frac{u_k - u_{k-1}}{h_{k-\frac{1}{2}}} + i\omega\mu_0 h_k \sigma_k u_k - h_k f_k \right) \end{aligned} \quad (\text{B.8})$$

and the adjoint-state equation of equation (B.3) is

$$\frac{\bar{\lambda}_{k+1}^b - \bar{\lambda}_k^b}{h_{k+\frac{1}{2}}} - \frac{\bar{\lambda}_k^b - \bar{\lambda}_{k-1}^b}{h_{k-\frac{1}{2}}} + i\omega\mu_0 h_k \sigma_k \bar{\lambda}_k^b = \overline{u_k - u_k^d}. \quad (\text{B.9})$$

It is easy to see that $\bar{\lambda}_k^a = h_k \bar{\lambda}_k^b$. The gradient of f is given by

$$\frac{\partial f}{\partial \sigma_k} = -\text{Re}(i\omega\mu_0 \bar{\lambda}_k^a u_k) = -\text{Re}(i\omega\mu_0 h_k \bar{\lambda}_k^b u_k). \quad (\text{B.10})$$

If now instead of using the adjoint-state system of the discrete finite-difference scheme, we discretize the continuous adjoint-state equation B.5 with the FDM used to discrete equation B.1, we obtain:

$$\frac{\bar{\lambda}_{k+1}^c - \bar{\lambda}_k^c}{h_k h_{k+\frac{1}{2}}} - \frac{\bar{\lambda}_k^c - \bar{\lambda}_{k-1}^c}{h_k h_{k-\frac{1}{2}}} + i\omega\mu_0\sigma_k \bar{\lambda}_k^c = \overline{u_k - u_k^d}; \quad (\text{B.11})$$

This leads to the approximated gradient

$$g_k^a = -\text{Re}(i\omega\mu_0 u_k \bar{\lambda}_k^c). \quad (\text{B.12})$$

Equations B.7 and B.11 correspond to two discretizations of the continuous adjoint-state equations. Therefore with a fine discretization, equation B.12 will give a good approximation of the discrete gradient since the approximation error in equation B.12 is roughly of the order of the discretisation of the direct problem. We, however, often use a discretization with stretching and as coarse as possible to reduce the memory

requirement and the computational time. We then prefer to implement the exact gradient of the discrete formulation to obtain a consistent code. Using the approximate gradient, equation B.12, instead of the exact gradient, equation B.10, may lead to a less efficient numerical optimization. However, we did not investigate this aspect.

References

- [1] Eidesmo T., Ellingsrud S., MacGregor L.M., Constable S., Sinha M.C., Johansen S. and Kong F.K. 2002 Seabed logging (SBL) a new method for remote and direct identification of hydrocarbon-filled layers in deep water areas *First Break* **20** 144–152.
- [2] Mac Gregor L., Andreis D., Tomlinson J. and Baker N 2006 Controlled source electromagnetic imaging on the Nuggets-1 reservoir *The Leading Edge* **25** 984-992.
- [3] Darnet M., Choo M.C.K., Plessix R.-E., Rosenquist M.L., Yip-Cheong K., Sims E. and Voon J.W.K. 2007 Detecting hydrocarbon reservoir from controlled source electromagnetic (CSEM) data in complex settings: application to deep water Sabah, Malaysia *Geophysics* **72** WA97–WA103.
- [4] Um E.S. and Alumbaugh D.L. 2007 On the physics of the marine controlled-source electromagnetic method *Geophysics* **72** WA13–WA26.
- [5] Habashy, T.M., Groom, R.W. and Spies, B.R. 1993 Beyond the Born and Rytov approximations: a nonlinear approach to electromagnetic scattering, *J. Geophys. Res.* **98** 1759–1775.
- [6] Mittet R., Maaø F., Aakervik O.M. and Ellingsrud S. 2005 A two-step approach to depth migration of low frequency electromagnetic data *Proc. of 67th EAGE Annual Meeting, Extended Abstract.*
- [7] Chavent G. 1989 A non-linear least-square theory for inverse problems *Inverse Methods in Action* ed. P. Sabatier (Springer).
- [8] Zhdanov M.S. 2002 Geophysical inverse theory and regularization problem (Elsevier).
- [9] Tarantola A. 2005 Inverse problem theory and model parameter estimation (SIAM).
- [10] Ward S.H. and Hohmann G.W. 1987 Electromagnetic theory for geophysical applications in *Electromagnetic Methods in Applied Geophysics – Theory* Vol. 1 M. N. Nabighian ed. (SEG) 131–311.
- [11] Mackie R.L., Madden T.R. and Wannamaker P.E. 1993 Three-dimensional magnetotelluric modeling using difference equations – Theory and comparison to integral equation solutions *Geophysics* **58** 215–226.
- [12] Wang T and Hohmann G.W. 1993 A finite-difference, time-domain solution for three-dimensional electromagnetic modeling *Geophysics* **58** 797–809.
- [13] Druskin V. and Knizhnerman L. 1994 Spectral approach to solving three-dimensional Maxwell’s diffusion equations in the time and frequency domain *Radio Sci.* **29** 937-953.
- [14] Newman G.A. and Alumbaugh D.L 1999 Electromagnetic modeling and inversion on massively parallel computers in *Three Dimensional Electromagnetics* M. Oristaglio and B. Spies eds (SEG).
- [15] Haber E., Ascher U., Aruliah D. and Oldenburg D. 2000 Fast simulation of 3D electromagnetic using potentials *J. Comp. Physics* **163** 150–171.
- [16] Haber E. and Ascher U. 2001 Fast finite volume simulation of 3D electro-magnetic problems with highly discontinuous coefficients *SIAM J. of Sci. Comp.* **22** 1943-1961.
- [17] Clemens M. and Weiland T. 2001 Discrete electromagnetism with the Finite Integration Technique *Progress in Electromagnetic Research (PIER)* **32** 65–87.
- [18] Yee K. 1966 Numerical solution of initial boundary value problems involving Maxwell’s equations in isotropic media *IEEE Transactions on Antennas and Propagation* **16** 302–307.
- [19] Van der Vorst H.A. 1992 BI-CGSTAB: a fast and smoothly converging variant of bi-CG for the solution of nonsymmetric linear systems *SIAM J. on Sci. and Stat. Comput.* **13** 631–644.
- [20] Gutknecht M.H. 1993 Variants of BiCGStab for matrices with complex spectrum *SIAM J. on Sci. and Stat. Comput.* **14** 1020–1033.

- [21] Aruliah D.A. and Ascher U.A. 2003 Multigrid preconditioning for Krylov methods for time-harmonic Maxwell's equations in 3D *SIAM J. on Sci. and Comp.* **24** 702–718.
- [22] Mulder W.A. 2006 A multigrid solver for 3D electromagnetic diffusion *Geophys. Prosp.* **54** 633–649.
- [23] Plessix R.-E., Darnet M. and Mulder W.A. 2007 An approach for 3D multi-source, multi-frequency CSEM modeling *Geophysics* **72** SM177 – SM184.
- [24] Ramm A.G. and Somersalo E. 1989 Electromagnetic inverse problem with surface measurements at low frequencies *Inverse Problems* **5** 1107–1116.
- [25] Mackie R.L. and Madden T.R 1993 Three-dimensional magnetotelluric inversion using conjugate gradients *Geophys. J. Int.* **115** 215-229.
- [26] Newman G.A. and Hoversten G. M. 2000 Solution strategies for two- and three-dimensional electromagnetic inverse problems *Inverse Problems* **16** 1357–1375.
- [27] Rodi W. and Mackie R.L. 2001 Nonlinear conjugate gradients algorithm for 2-D magnetotelluric inversion *Geophysics* **66** 174-187.
- [28] Haber E., Ascher U., Aruliah D. and Oldenburg D. 2004 Inversion of 3D electromagnetic data in frequency and time domain using an inexact all-at-once approach *Geophysics* **69** 1216–1228.
- [29] Carazzone J.J., Burtz K.E., Green K.E. and Pavlov D.A. 2005 Three dimensional imaging of marine CSEM data *Proc. of 75th SEG Annual Meeting, Expanded Abstract* **24** 575–578.
- [30] Newman G.A. and Boggs P. T. 2004 Solution accelerators for large-scale three-dimensional electromagnetic inverse problems *Inverse Problems* **20** S151-S170.
- [31] Haber E. 2005 Quasi-Newton methods for large-scale electromagnetic inverse problems *Inverse Problems* **21** 305–223.
- [32] Gilbert J.C. and Lemaréchal C. 1989 Some numerical experiments with variable storage quasi-Newton algorithms *Mathematical Programming* **45** 407–436.
- [33] Byrd R.H., Nocedal J. and Schnabel, R.B. 1994 Representation of quasi-Newton matrices and their use in limited memory methods *Mathematical Programming* **63** 129–156.
- [34] Byrd R.H., Lu P. and Nocedal J. 1995 A limited memory algorithm for bounded constrained optimization, *SIAM J. on Sci. and Stat. Comp.* **16** 1190-1208.
- [35] Chavent G. and Lemonnier P. 1974 Identification de la Non-Linéarité d'une équation parabolique quasilinéaire *Applied Mathematics and Optimization* **1** 121–162.
- [36] Plessix R.-E. 2006 A review of the adjoint-state method for computing the gradient of a functional with geophysical applications *Geophys. J. Int.* **167** 495–503.
- [37] Monk P. and Süli E. 1994 A convergence analysis of Yee's scheme on nonuniform grids *SIAM J. of Num. Anal.* **31** 393–412.
- [38] Gill P.E., Murray W. and Wright M.H. 1981 *Practical Optimization* (Academic Press, London).
- [39] Newman G.A. and Alumbaugh D.L 2000 Three-dimensional magnetotelluric inversion using non-linear conjugate gradients *Geophys. J. Int.* **140** 410–424.
- [40] Nocedal, J. 1980 Updating quasi-Newton matrices with limited storage *Mathematics of computation* **35** 773–782.
- [41] Plessix R.E. and Mulder W.A. 2004 Frequency-domain finite-difference amplitude-preserving migration *Geophys. J. Int.* **157** 975–987.