Near Surface Geoscience 2012

A18

Permittivity and Conductivity Reconstruction by Full Waveform Inversion of GPR Data using the L-BFGS-B Algorithm

F. Lavoué* (Institut des Sciences de la Terre,), R. Brossier (Institut des Sciences de la Terre), S. Garambois (Institut des Sciences de la Terre) & J. Virieux (Institut des Sciences de la Terre)

SUMMARY

Full waveform inversion (FWI) of ground-penetrating radar data is an emerging technique for quantitative imaging of the near surface, mainly through the estimation of the dielectric permittivity (ε) and of the electric conductivity (σ). Recent studies already succeeded to provide high resolution cross-hole images by FWI using conjugate gradient algorithms. In this study, we present a frequency-domain FWI algorithm based on the L-BFGS-B optimization which takes into account the Hessian influence in the steepest-descent direction correcting for dimensionalities between parameters. We discuss the impact of the parametrization for the simultaneous reconstruction of ε and σ, showing that a robust criterion is provided by the ratio of the gradient norms in the directions of ε and σ: the relative amplitudes of the gradients of the chosen parameters greatly impact the conditioning of the inverse problem. We show that the sensitivity of the cost function to the selected parameters needs to be taken into account in the re-parametrization. An illustration is provided using an already published benchmark. It demonstrates the great efficiency of the L-BFGS-B optimization method to deal with non-linearities of the inverse problem.

Introduction

Ground-penetrating radar (GPR) is a non-invasive prospecting technique based on the propagation of electromagnetic waves. The GPR principle is very close to seismic methods and therefore take large benefits from seismic interpretation techniques, so that it provides today accurate qualitative images of the subsurface. However, the development of a quantitative imagery that would estimate the electromagnetic proprieties of the investigated medium (electric conductivity σ [S/m] and dielectric permittivity ε [F/m], mainly) remains a critical issue to derive an accurate interpretation of natural structures. Recently, efforts have been made towards quantitative GPR imagery using ray tomography, AVO studies and full-waveform inversion (FWI). Already well developed in seismics (see Virieux and Operto, 2009, for a review), FWI has been recently applied in GPR for the imaging of 2D cross-hole sections by Meles et al. (2010) in the time-domain, and by Ellefsen et al. (2011) in the frequency-domain. Most of these algorithms are based on conjugate gradient methods to minimize the difference between observed and synthetic data. Here, we propose a 2D frequency-domain FWI method based on the limited Broyden-Fletcher-Goldfarb-Shanno bounded algorithm (L-BFGS-B, Byrd et al., 1995). A major feature of L-BFGS is the estimation of a cost effective influence of the Hessian on the current steepest-descent direction from gradients at previous iterations. This non-diagonal approximate correction of the direction allows to consider high contrasts, improves focusing, partially corrects the descent direction from effects due to limited aperture illumination and frequency bandwidth and respects dimensions of the different parameter values for multi-parameter inversion (Brossier et al., 2009). Furthermore, considering the bounded version of the algorithm (L-BFGS-B) is of great interest for GPR imagery, since physical limits are often encountered ($\varepsilon_r = 1$ and $\sigma = 0$ S/m in the air). After presenting the algorithm, we show its efficiency on a synthetic case with perfect illumination, inspired from Meles et al. (2011). We will particularly underline the influence of the parametrization for simultaneous reconstruction of ε and σ.

General theory

The forward problem is solved using an optimized frequency-domain finite-difference scheme, originally developed for seismic waves (Hustedt et al., 2004). This modeling tool is adapted for the 2D electromagnetic problem, thanks to the mathematical analogy between both systems of equations (e.g. Carcione and Robinson, 2002). Diffusive losses are easily considered in this frequency approach through the complex relative permittivity $\varepsilon_r^* = \varepsilon_r + i\sigma/(\varepsilon_o\omega)$, where $\varepsilon_o \simeq 8.85.10^{-12}$ F/m is the permittivity of vacuum and $\varepsilon_r = \varepsilon/\varepsilon_o$. In the following, we are interested in reconstructing images of ε_r and σ from measurement of the $\mathbf{E}_y$ component of the electric field [V/m] in the Transverse Electric mode.

We use the classical misfit function $C(\mathbf{m})$ defined with respect to the model $\mathbf{m}$ using the ℓ_2 norm as $C(\mathbf{m}) = \{\Delta\mathbf{d}^\dagger\Delta\mathbf{d}\}/2$, where $\Delta\mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}(\mathbf{m})$ is the difference between observed ($\mathbf{d}_{obs}$) and synthetic ($\mathbf{d}(\mathbf{m})$) data. The transpose (^{T}) conjugate (*) operator is denoted by †. Note that we do not use any regularization term since we aim to discuss effects of the parametrization on the conditioning of the problem. The gradient of the cost function is computed using the adjoint state method (Plessix, 2006):

$$\mathbf{G}(m_i) = -\mathscr{R}e\left\{\sum_{\omega_i=1}^{N_\omega}\sum_{s=1}^{N_S}\mathbf{E}_y\left(\frac{\partial\mathbf{A}}{\partial m_i}\right)\mathbf{A}^{-1^\dagger}\Delta\mathbf{d}\right\}, \tag{1}$$

where $\mathbf{A}$ is the impedance matrix, resulting from the discretization of the forward problem. Note the sum over the number of sources N_S and the number of used frequencies N_ω, which have implications on the parametrization that we discuss later.

We focus our attention on the diffraction matrix $(\partial\mathbf{A}/\partial m_i)$ to characterize the gradient of the cost function with respect to parameters ε_r and σ. The diffraction matrix of these parameters can be expressed as

$$\partial_{\varepsilon_r}\mathbf{A} = -\omega^2\varepsilon_o, \quad \text{and} \quad \partial_\sigma\mathbf{A} = -i\omega. \tag{2}$$

Unfortunately, quantities ε_r and σ neither have the same units nor the same order of magnitude ($\varepsilon_r \in$

$[1,81]$ whereas $\sigma \in [0,1]$ S/m, roughly). The gradient using the expressions of eq. (2) would, therefore, keep the footprint of parameters units and amplitudes and could not be used directly to define a consistent optimization scheme. As a consequence, an appropriate scaling, or re-parametrization, is required.

Parametrization based on the norm of the gradient

An ideal parametrization should balance equally well the impact of the parameters on the cost function and should account for their dynamics. Ellefsen et al. (2011) invert the logarithm of the complex slowness for instance. An other example is provided by Meles et al. (2010), who implicitly change their parametrization by defining two descent step lengths, for ε_r and σ, respectively. This amounts to scale the gradients differently according to values of ε_r and σ. In this study, we promote a linear re-parametrization on the reconstructed parameters, that directly and consistently impacts the whole optimization process. This strategy guarantees the convergence properties of L-BFGS-B with a single step length (respect of the Wolfe conditions in the line-search).

An arbitrary rescaling is applied to the parameters in such a way that the gradient has approximatively the same norm in the direction of ε_r ($|\mathbf{G}_{|\varepsilon_r}|$) as in σ ($|\mathbf{G}_{|\sigma}|$). This approach is tested on a benchmark inspired from Meles et al. (2011), which consists in two cross-shaped anomalies with $\varepsilon_{r_1} = 8$, $\varepsilon_{r_2} = 1$, $\sigma_1 = 10$ mS/m and $\sigma_2 = 0.1$ mS/m, in a background with $\varepsilon_r = 4$ and $\sigma = 3$ mS/m. As Meles et al. (2011), we consider a perfect illumination with 11 sources and 31 receivers on each edge of the medium. The initial model is an homogeneous medium with the values of the background. According to Meles et al. (2011), the two crosses are close enough to induce multiple scattering and to challenge the inversion.

Computing the gradient of the standard parameters at the first iteration (as defined by eq. (1) and (2)) allows to measure the ratio $|\mathbf{G}_{|\sigma}|/|\mathbf{G}_{|\varepsilon_r}| = R$. We can define a set of new reconstructed parameters defined as $\tilde{\varepsilon}_r = \varepsilon_r$ and $\tilde{\sigma} = R\sigma$. We have therefore $\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r} = \mathbf{G}_{|\varepsilon_r}$ and $\tilde{\mathbf{G}}_{|\tilde{\sigma}} = \partial_{\tilde{\sigma}} C = \partial_\sigma C . \partial_{\tilde{\sigma}} \sigma = \mathbf{G}_{|\sigma}/R$, and we expect to get $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}| \simeq 1$.

We apply this approach to the numerical benchmark, inverting simultaneously for 7 frequencies from 50 up to 300 MHz, and obtain an acceptable image of ε_r (see figure 1a) but a noisy reconstruction of σ (fig. 1b). Choosing a ratio such that $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}| = 1$ is not appropriate, as the sensitivities of permittivity and conductivity are quite different in the cost function.

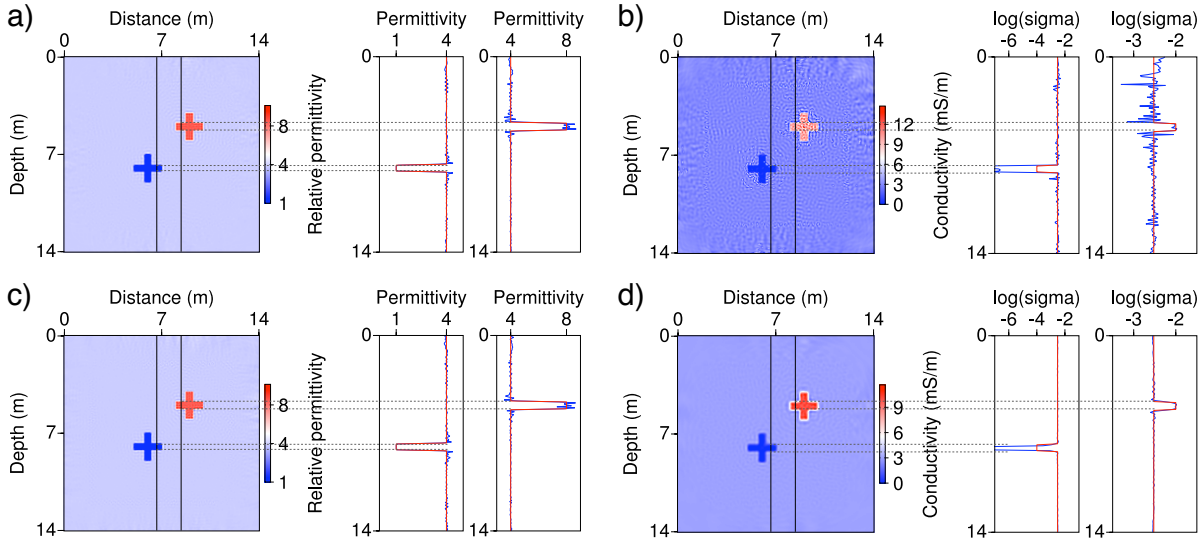

Figure 1 *Inversion with 7 simultaneous frequencies (50, 70, 100, 150, 200, 250, and 300 MHz), using a parametrization such that* $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}| \simeq 1$ *(a, b) and such that* $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}| \simeq 0.1$ *(c, d).*

Figure 2 presents a grid analysis of the cost function for a simplified case: we modify the benchmark to

a two-parameters problem in which only the values of ε_r and σ of the positive anomaly (with $\varepsilon_{r_1} = 8$ and $\sigma_1 = 10$ mS/m) are allowed to vary. This first cross is considered homogeneous and the background (including the second cross) is kept to its true value, such that the problem has only two degrees of freedom: ε_{r_1} and σ_1. The cost function has been evaluated for ε_{r_1} ranging in [1,81] and σ_1 in [0,1] S/m. Figure 2a exhibits a strong asymmetry of the valley around the global minimum, indicating that the cost function is much more sensitive to changes in ε_r than in σ in the vicinity of the solution (fig. 2b).

Therefore, we change our parametrization such that $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}| \simeq 0.1$: this scaling takes much better into account the natural sensitivity of the cost function (in other words, the data) to each parameter. The image of σ has been strongly improved with this second parametrization (see figure 1c, d). Changing the ratio from 1 to 0.1 acts as a preconditioning on the ill-posed optimization problem. This indicates that the inversion is mainly driven by the permittivity, which dominates the sensitivity of the cost function.

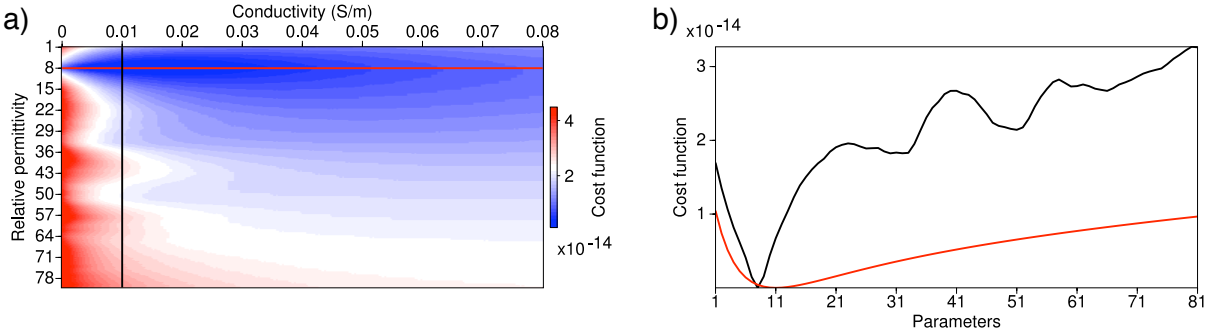

Figure 2 *Grid analysis of the cost function for a simplified two-parameters problem. a) 2D grid analysis. b) 1D-sections along ε_r for $\sigma = \sigma^{true} = 10$ mS/m (black line) and along σ for $\varepsilon_r = \varepsilon_r^{true} = 8$ (red line).*

Discussion

Our study shows that a re-parametrization, based on the observed ratio $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}|$ and on a weighting derived from the sensitivity of the cost function, allows to tackle simultaneous reconstruction of permittivity and conductivity. The weighting can be considered as a tuning term which seems arbitrary and difficult to determine as the grid analysis performed here cannot be achieved in all cases. Nevertheless, in most cases, the cost function is more sensitive to ε_r than to σ, with a ratio of about 0.01 to 0.1. Furthermore, a quality control can be applied on this weighting as it acts as a preconditioning: if the conductivity drives the inversion too strongly, noisy images of σ will be obtained. Conversely, if the weight of conductivity is too weak, it will provide smoother images but will damage the resolution of σ.

We can notice in figure 1 that the magnitude of σ in the negative anomaly is not recovered exactly, but we reach a rough value of 0 S/m corresponding to the bound. To improve this result, instead of inverting for the 7 frequencies simultaneously, we adopt a hierarchical strategy consisting in gathering some frequencies within groups that are inverted sequentially (Brossier et al., 2009). For each group, the starting model results from the inversion of the previous group and a re-parametrization is applied.

Inverting for the frequency groups [50, 70, 100], [150, 200] and [250, 300] MHz, we effectively retrieve the exact value of 0.1 mS/m in the anomaly. Since the impact of σ on the data is frequency-dependent (see eq. (2) for the effect on the diffraction matrix), considering groups of narrow frequencies allows to better take benefit of high frequencies than inverting simultaneously all the frequencies. In other words, the simultaneous inversion of all frequencies decreases the weight of the high-frequency components with respect to the low frequency components. Figure 3 illustrates this effect by presenting the ratio $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}|$ during the sequential inversion of the same 7 frequencies, using a constant parametrization designed such that $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}| \simeq 0.1$ at 150 MHz. The ratio $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|/|\tilde{\mathbf{G}}_{|\tilde{\varepsilon}_r}|$ clearly decreases with respect to frequency. The panels showing the reconstructed models at the end of each mono-frequency inversion indicates that σ is no more updated in the last three groups, due to the too weak weight of $|\tilde{\mathbf{G}}_{|\tilde{\sigma}}|$.

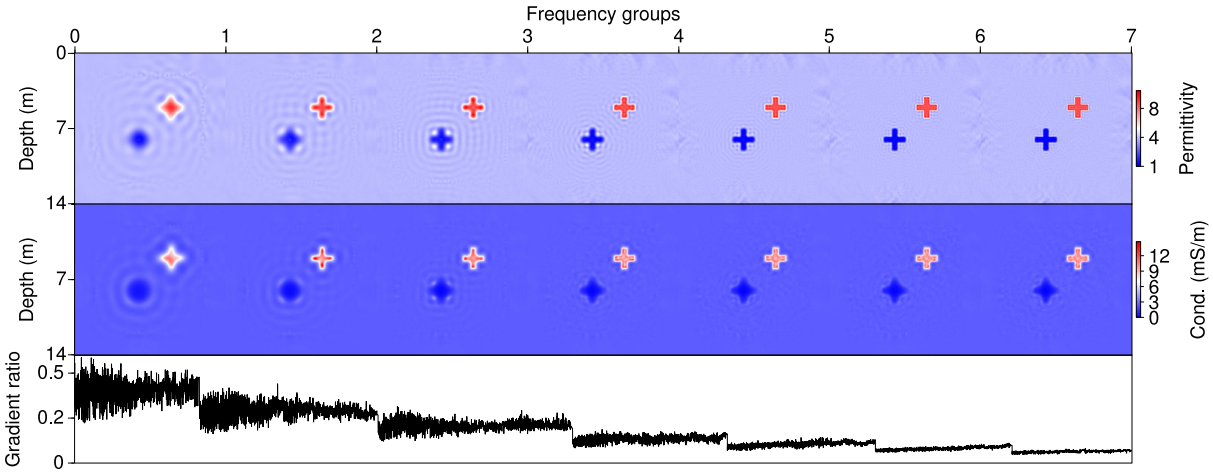

Figure 3 *Evolution of the reconstructed models and of the gradient ratio* $|\tilde{\mathbf{G}}_{|_{\tilde{\sigma}}}|/|\tilde{\mathbf{G}}_{|_{\tilde{\varepsilon}_r}}|$ *vs. frequency groups when the 7 frequencies are sequentially inverted using the same parametrization.*

Conclusions

We present a FWI algorithm based on a quasi-newton L-BFGS-B optimization for the simultaneous reconstruction of the parameters ε_r and σ. This multi-parameter inversion requires the design of a correct parametrization for balancing the respective footprint of parameters on the cost function. A robust criterion is provided by the measurement of the ratio between the norms of the gradient for ε_r and σ at the first iteration. This ratio must be weighted according to the sensitivity of the cost function with respect to the parameters in order to ensure an improved conditioning of the optimization. This criterion is believed to be a general result, which can be applied as a guiding rule to any type of parametrization. However, its robustness when noise is present in the data is still to be established. The algorithm has been tested on a known benchmark, showing the great efficiency of L-BFGS-B for dealing with non-linearities during optimization. Further work should investigate its capability to accurately retrieve high contrasted anomalies and to manage partial illumination in cross-hole or surface-to-surface acquisitions.

Acknowledgments

This work was performed using HPC resources from GENCI-CINES (Grant 2011-046091]) and from the CIMENT computing center (Univ. Grenoble). It was partially funded by the SEISCOPE consortium.

References

Brossier, R., Operto, S. and Virieux, J. [2009] Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, **74**(6), WCC63–WCC76, doi:10.1190/1.3215771.

Byrd, R., Lu, P. and Nocedal, J. [1995] A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific and Statistical Computing*, **16**, 1190–1208.

Carcione, J.M. and Robinson, E.M. [2002] On the acoustic-electromagnetic analogy for the reflection-refraction problem. *Studia Geophysica et Geodaetica*, **46**, 321–346.

Ellefsen, K.J., Mazzella, A.T., Horton, R.J. and McKenna, J.R. [2011] Phase and amplitude inversion of crosswell radar data. *Geophysics*, **76**(3), J1–J12.

Hustedt, B., Operto, S. and Virieux, J. [2004] Mixed-grid and staggered-grid finite difference methods for frequency domain acoustic wave modelling. *Geophysical Journal International*, **157**, 1269–1296.

Meles, G.A., Greenhalgh, S., van der Kruk, J., Green, A.G. and Maurer, H. [2011] Taming the non-linearity problem in GPR full-waveform inversion for high contrast media. *Journal of Applied Geophysics*, **73**, 174–186.

Meles, G.A., der Kruk, J.V., Greenhalgh, S.A., Ernst, J.R., Maurer, H. and Green, A.G. [2010] A New Vector Waveform Inversion Algorithm for Simultaneous Updating of Conductivity and Permittivity Parameters From Combination Crosshole/Borehole-to-Surface GPR Data. *IEEE Transactions on Geoscience and Remote Sensing*, **48**, 3391–3407.

Plessix, R.E. [2006] A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, **167**(2), 495–503.

Virieux, J. and Operto, S. [2009] An overview of full waveform inversion in exploration geophysics. *Geophysics*, **74**(6), WCC127–WCC152.