Full-waveform inversion of GPR data acquired between boreholes in Rustrel carbonates

Hugo Pinard[1,a], Stéphane Garambois[1], Ludovic Métivier[1], Michel Dietrich[1], Guy Sénéchal[2], and Dominique Rousset[2]

[1] Université Grenoble Alpes, ISTerre, CNRS, 38000 Grenoble, France
[2] LFCR, Université de Pau et des Pays de l'Adour, 64000 Pau, France

Abstract. Full waveform inversion (FWI) of seismic or Ground Penetrating Radar data provides high-resolution quantitative images of the constitutive parameters of the rock/soil which control seismic/GPR wave propagation. We developed a 2D inversion tool in the frequency domain adapted to the multi-parameter physics controlling GPR propagation in isotropic non dispersive media, i.e. dielectric permittivity and electrical conductivity. This inversion engine was previously tested using synthetic 2D data to mitigate the trade-off between the two parameter classes. In this paper, we present the required processing techniques and first inversion results obtained on a real GPR dataset acquired in carbonates with a cross-hole configuration. The presence of the 2 m diameter underground gallery at depth constitutes a nice target to test the robustness, efficiency and resolution of the inversion in such high-contrasts context. Starting from a time tomographic image for the dielectric permittivity and from a homogeneous conductivity, we show that FWI is efficient to retrieve high resolution images of dielectric permittivity but struggles with electrical conductivity. As a quality control, we compare real and synthetic radargrams computed from the tomography and final images, showing the efficiency of the process to reconstruct some events but also underlying some issues, particularly on large incidence angles amplitude traces.

1. Introduction

The knowledge of the subsurface is crucial for a wide range of scientific issues such as environmental, archaeological, hydrological, geological and civil engineering. In this frame, Ground Penetrating Radar (GPR) plays a key role thanks to its non-invasive and high resolution features. Multioffset measurements may provide quantitative information of the subsurface through velocity analysis or inverse processes like traveltime tomography or full waveform inversion (FWI). With traveltime tomography, only a small fraction of the recorded data is involved in the inversion providing a smooth quantitative image of the dielectric permittivity with a limited resolution. In this context, FWI, which takes benefit from the whole recorded signal, should improve the image resolution.

FWI of cross-hole GPR datasets has shown its efficiency on synthetic and real [1] cases. Here, we perform FWI in carbonate environment with the favorable cross-hole configuration in order to evaluate our multi-parameter inversion algorithm. Based on previous

---

[a] Corresponding author: `hugo.pinard@univ-grenoble-alpes.fr`

synthetic tests, we use the l-BFGS algorithm which is well known in seismic to reconstruct simultaneously subsurface parameters.

Here, we briefly explain the concepts underlying the FWI with the description of the forward and inverse problems. We then describe the pre-processing methodologies used to estimate the source wavelet and to convert real data acquired in a 3D environment to 2D data managed by the code. Finally, we present our preliminary FWI result which contains many artifacts. Despite these artifacts, which emphasize the current failure of the inversion, we show the first benefits of the method.

2. Theoretical basics on forward and inverse problems

In 2D, the introduction of the constitutive properties in the Maxwell-Faraday and Maxwell-Ampère equations gives two distinct systems: the transverse electric (TE) and the transverse magnetic (TM) modes. Here, we consider only the TM one with vertical electric sources, as used in boreholes. The system of three first-order equations is recast into one second-order equation and two first-order equations. This transformation and acoustic/electromagnetic analogy [2] allow to simulate electromagnetic wave propagation using finite-difference frequency-domain schemes initially developed for seismic modeling in visco-acoustic approximation [3]. This formulation leads to the system $A(m)H_y = s$ where the discrete wave operator $A(m)$ contains the subsurface parameters m while H_y and s are respectively the magnetic wavefield and the discretized source terms. This linear system is solved using a LU factorization through a sparse direct solver [4]. The electrical fields E_x and E_z, associated to GPR antennas, are deduced by finite-difference of H_y following the two first-order equations.

This frequency domain formulation allows to solve TM equations at a reasonable computational cost incorporating easily the diffusive attenuation through the complex effective permittivity $\varepsilon_e = \varepsilon + i\sigma/\omega$ with the dielectric permittivity ε [F/m], the electrical conductivity σ [S/m], the angular frequency ω [rad/s] and the pure imaginary number (using a conventional time-harmonic dependency in $e^{-i\omega t}$).

The formalism of FWI has been introduced in the seismic context by [5] and [6] in the time domain and developed later by [7] in the frequency domain. The target of the FWI is to find the distribution of parameters m which minimizes the real misfit function $f(m)$. This quantity is defined as the distance between the observed data $d_{obs,\omega_j,s}$ and the calculated data $d_{cal,\omega_j,s}(m)$, namely the residuals $\Delta d_{\omega_j,s}(m) = d_{obs,\omega_j,s} - d_{cal,\omega_j,s}(m)$, through the ℓ_2 norm:

$$f(m) = \frac{1}{2} \sum_{j=1}^{N_\omega} \sum_{s=1}^{N_s} \Delta d_{\omega_j,s}(m)^\dagger \Delta d_{\omega_j,s}(m) \tag{1}$$

where the symbol † denotes the transpose conjugate operator while the total number of sources and discrete frequencies are denoted by N_s and N_ω respectively. In our particular case, observed and calculated data correspond to the vertical components of the electric field (Ez) at receiver positions. The FWI procedure iteratively updates the model of the subsurface through the formula:

$$m_{k+1} = m_k + \alpha_k \Delta m_k, \tag{2}$$

where α_k is the linesearch parameter satisfying the Wolfe conditions and Δm_k is the descent direction solution of the Newton equation $H_k \Delta m_k = -\nabla f_k$. Among the different methods to solve the Newton equation, we choose the l-BFGS one for its efficiency and accuracy. With this method, the inverse of the Hessian operator $H(m_k)^{-1}$ is approximated through finite-difference of l previous gradients $\nabla f(m_{k-1})...\nabla f(m_{k-l})$. Thus, the algorithm only requires to

compute the gradients of each parameter. These quantities are obtained very efficiently thanks to the first-order adjoint-state method ([8] for a review). In the following application, we also introduce the Tikhonov regularization to smooth as much as possible the model update.

3. Application

Data were acquired in an Urgonian karstic limestone massif located near the LSBB. This underground structure goes through the mountain (Fig. 1a). One of its gallery, located close to the exit (S), is relatively shallow and can thus be investigated from boreholes. Among the different boreholes drilled to study carbonate properties, two of them surrounded the gallery (Fig. 1b). This configuration allowed to perform measurements in cross-hole configuration.

Two 100 MHz antennas have been used to record the dataset between these two 50 m deep boreholes separated by 18 m. 44 sources are evenly distributed (1 m interval) in borehole F5 while 200 traces have been recorded every 25 cm in borehole F6. Examples of radargram are displayed in Fig. 1c and d for two different source positions. The first one shows diffractions only between traces 70 and 100 while the second one is more complex, exhibiting three distinct hyperbolae with apex at traces 120, 135 and 145.

All first arrivals have been picked and inverted using a SIRT algorithm to get a smooth image of the velocity and so of the dielectric permittivity (Fig. 2a). The 2 m diameter underground gallery at depth 34 m is visible but appears quite spread compared to the reality where sharp contrasts exist at the interfaces air/concrete/carbonates. We expect that FWI provides higher resolution images in particular of the gallery.

Before applying FWI, a pre-processing of the data is required. We first remove the average value of each trace and apply a Butterworth filter to cut low and high frequency noises. Since our forward modeling engine is 2D and the observed data ($d_{obs}^{3D}(t)$) are 3D, we apply a 3D-to-2D conversion to get the 2D observed data ($d_{obs}^{2D}(t)$). The conversion, given here for one trace, is performed through the formula,

$$d_{obs}^{2D}(t) = \sqrt{\frac{2\pi v(t)^2 t}{\omega}} e^{i\pi/4} d_{obs}^{3D}(t) \tag{3}$$

with t denotes the time, w the angular frequency, i imaginary number and $v(t)$ the wave velocity which is assumed to be homogeneous (in the formula) and equals to 8 cm/ns following the traveltime tomography result. One example of this conversion is displayed in Fig. 1e where changes in phase and amplitude can be clearly seen. Then, the dataset is normalized by shot-gather to be sure that none of these shot-gathers dominates the misfit function. Note that frequency content of the signal is extracted and displayed in Fig. 1f. In general, all frequencies between 50 to 150 MHz have a strong amplitude so that this frequency band can be used to perform the FWI.

The last step of the pre-processing consists in estimating the source wavelet. In the frequency-domain, this estimation, detailed here for one frequency ω_j and one source s, is classically realized through the linear inversion proposed by [9],

$$S_{\omega_j,s} = \frac{g_{\omega_j,s}^{\dagger} \, d_{obs,\omega_j,s}^{2D}}{g_{\omega_j,s}^{\dagger} \, g_{\omega_j,s}}, \tag{4}$$

where $S_{\omega_j,s}$ is the complex value of the source wavelet at frequency ω_j while $g_{\omega_j,s}$ and $d_{obs,\omega_j,s}^{2D}$ denote respectively the incident Green's functions recorded at receiver positions and the observed data in the frequency domain. Here, each source s is supposed independent from each other so that one source estimation is done per shot-gather. After computations of $S_{\omega_j,s}$ for all frequencies and inverse Fourier transform, the resulting source wavelet for one
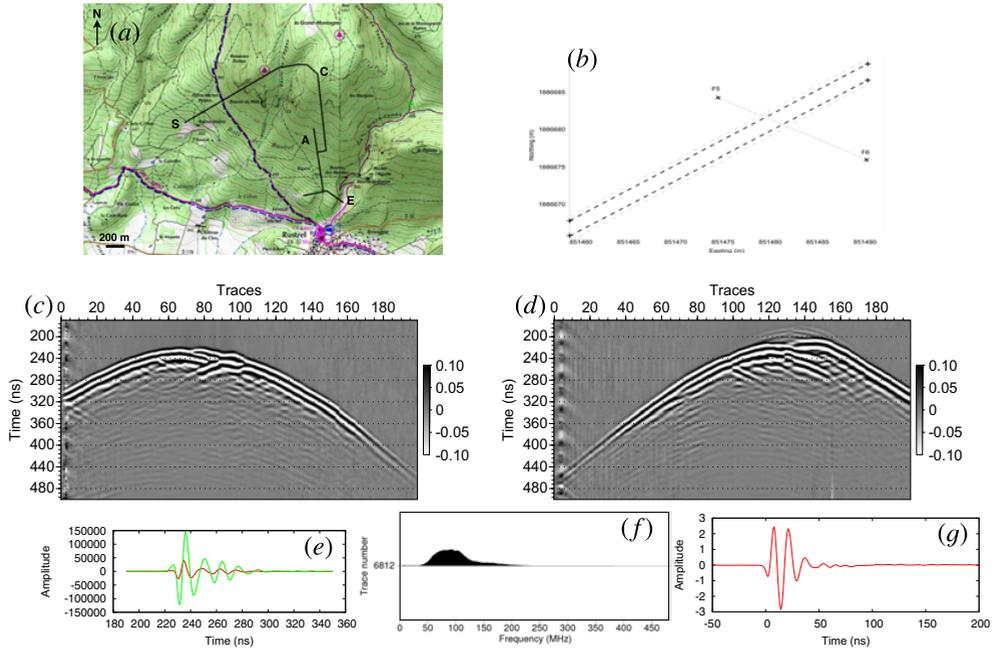
Figure 1. (a) Location of LSBB galleries (black line). E denotes the main entrance, A the anti blast gallery, C the shielded LSBB capsule and S the exit. (b) Zoom close to the exit with the locations of the two boreholes of interest and the gallery. Observed radargrams for source positions at depth (c) 18 m and (d) 34 m. (e) 3D-to-2D conversion for one trace the 3D (red) and 2D (green) traces. (f) Frequency content of one trace. (g) Source wavelet estimated before the inversion for one shot-gather.

source located at depth 17 m is displayed in Fig. 1g. The wavelet may be compared with a Ricker wavelet (second-order derivative of a Gaussian distribution) even if side rebounds are currently quite strong. Even if the wavelet can be updated regularly during the inversion, we choose to estimate it once at the beginning and keep fixed after.

The 51.5 m by 18.5 m model is discretized using a grid interval of 0.05 m to ensure that electromagnetic modeling satisfies the sampling criterion of four/five grid points per minimum wavelength. The initial model for the permittivity is provided by the traveltime tomography while the conductivity one is taken homogeneous (Fig. 2a and d).

The discrete frequencies strategy used to invert data follows the strategy developed by [10] and the frequency content of the dataset (Fig. 1f). One example of this strategy is presented in Table 1. Frequencies of a given group are used to perform one nonlinear inversion and associated results are used as initial models for the next group. Starting by the lower frequencies to reconstruct the large structures of the medium, higher frequencies are progressively introduced to improve the resolution of the model. Note that the frequency sampling ensures a continuous sampling in the wavenumber space. In this paper, we restrict results to the first group because we need to improve some points before using the high frequency content.

As discussed by [11], we invert relative parameters instead of parameter themselves so that they are defined as $\varepsilon_r = \varepsilon/\varepsilon_0$ where $\varepsilon_0 = 8.85 \cdot 10^{-12} \ Fm^{-1}$ (vacuum value) and $\sigma_r = \sigma/\sigma_0$ where σ_0 has to be chosen accurately. If this parameter is too high, only the
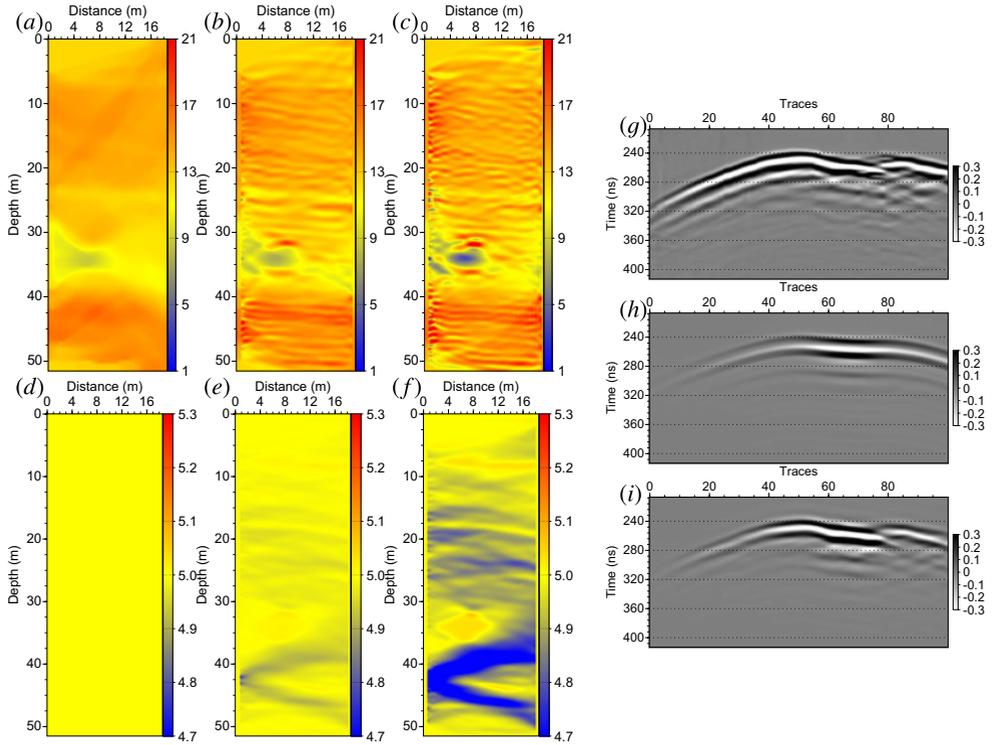
Figure 2. (a–c) Relative permittivity after 0, 15 and 30 iterations; (d–f) conductivity [mS/m] after 0, 15 and 30 iterations. (g) observed data, (h) initial data and (i) final data for 100 receivers associated to one source located at depth 39 m.

Table 1. Frequency strategy for the inversion (MHz).

Group 1	50	60	70
Group 2	60	70	80
Group 3	70	80	100
Group 4	80	100	120
Group 5	100	120	150

conductivity is updated and vice versa for the permittivity. In this paper, $\sigma_0 = 1.0 \ 10^{-4} \ S/m$ to slightly promote the relative permittivity update which stabilizes the inversion. Areas around sources and receivers have been frozen to prevent the model being updated only at these positions. We decided to perform 30 nonlinear iterations. The normalized misfit functions with respect to the initial value are 0.8 and 0.78 respectively at the 15th and 30th iterations. This small decrease is linked to difficulties for the algorithm to remove the imprint of sources and receivers despite the presence of the frozen zone (Fig. 2). At the end, sources and receivers, through their vertical dipolar injection points, create horizontal artifacts in the model. These latter, which present a higher wavenumber content in the permittivity model, still require a deeper analysis to better understand their origins and remove them.

Despite these artifacts, we obtain an increase of the resolution of the gallery focusing the permittivity and providing a first image of conductivity. In Fig. 2, the limit of the gallery becomes progressively sharper while the final relative permittivity value in the gallery is

closer to the expected one ("1" for the void). For the conductivity model, the gallery appears more conductive than the geological medium and the presence of water-saturated layers around 38 m depths is not detected. These results can be explained either by the presence of iron in the concrete of the gallery (even if its presence is not confirmed) or more probably by an initial model of conductivity too simple and/or too far from the true one.

Because the FWI is based-on data fitting procedure, radargrams are shown in Fig. 2g, h and i. Comparing observed and initial data, several differences can be underlined. Before trace 45, a clear mismatch in amplitude is visible while after this trace, the initial data do not reproduce the complex structure of the observed data and specially the discontinuity around trace 77. The analysis of the final radargram shows that the mismatch in amplitude is still present meaning that the algorithm did not update the model to reduce this difference. Actually, the algorithm concentrates its efforts on the complex arrivals and the discontinuity which are well retrieved at the end of the inversion. This final radargram, obtained with only one group of frequencies, emphasizes the potential of the FWI.

4. Conclusion

We briefly propose the formulation of a 2D frequency-domain FWI for GPR datasets in TM configuration to reconstruct simultaneously dielectric permittivity and electrical conductivity. Then, we present our first results using a dataset recorded near the LSBB. These results are promising with a focalization of the gallery in permittivity and a first image of conductivity. This work will be continued with a better handling of the source estimation (at each nonlinear iteration), the consideration of all groups of frequencies and other initial conductivity models. These different points should improve the quality and the resolution of the final reconstructions and, we hope, allow the detection of multi-scale heterogeneities characterizing the carbonate properties of this massif. Final results will be compared to well logs, electrical tomography and checked against ground truth (for the gallery location). At middle/long term, new acquisitions will be performed and results will be used to follow the evolution of the massif properties over time.

References

[1] A. Klotzsche, J. van der Kruk, G.A. Meles, H. Vereecken, Geophysics, **77**, H57 (2012).

[2] J.M. Carcione and F. Cavallini, Wave motion, **21**, 149 (1995).

[3] B. Hustedt, S. Operto, and J. Virieux, Geophysical Journal International, **157**, 1269 (2004).

[4] MUMPS-team. *MUMPS – Multifrontal Massively Parallel Solver users' guide – version 4.9.2 (November 5, 2009)*. ENSEEIHT-ENS Lyon, `http://www.enseeiht.fr/apo/MUMPS/` or `http://graal.ens-lyon.fr/MUMPS` February (2009).

[5] P. Lailly, *The seismic problem as a sequence of before-stack migrations*, in *Conference on Inverse Scattering: Theory and Applications*, edited by J. Bednar, (SIAM, Philadelphia, 1983).

[6]   A. Tarantola, Geophysics, **49**, 1259 (1984).
[7]   R.G. Pratt and M.H. Worthington, Geophysical Prospecting, **38**, 287 (1990).
[8]   R.E. Plessix, Geophysical Journal International, **167**, 495 (2006).
[9]   R.G. Pratt, Geophysics, **64**, 888 (1999).
[10] L. Sirgue, R.G. Pratt, Geophysics, **69**, 231 (2004).
[11] F. Lavoué, R. Brossier, L. Métivier, S. Garambois, J. Virieux, Geophysical Journal International, **197**, 248 (2014).