

Joint inversion of full-waveform GPR and ER data. Part 2: enhancing low frequencies with the envelope transform and cross-gradients

Diego Domenzain*, John Bradford†, Jodi Mead*

ABSTRACT

Recovering material properties of the subsurface using ground penetrating radar (GPR) data of finite bandwidth with missing low frequencies, and in the presence of strong attenuation is a challenging problem. We propose three non-linear inverse methods for recovering electrical conductivity and permittivity of the subsurface by joining GPR multi-offset and electrical resistivity (ER) data acquired at the surface. All methods use ER data to constrain the low spatial-frequency of the conductivity solution. The first method uses the envelope of the GPR data to exploit low frequency content in full-waveform inversion and does not assume structural similarities of material properties. The second method uses cross-gradients to manage weak amplitudes in the GPR data by assuming structural similarities between permittivity and conductivity. The third method uses both the envelope of the GPR data and the cross-gradient of the model parameters. By joining ER and GPR data, exploiting low frequency content in the GPR data, and assuming structural similarities between electrical permittivity and conductivity we are able to recover subsurface parameters in regions where the GPR data has a signal-to-noise ratio close to one.

INTRODUCTION

Electrical properties in the subsurface such as electrical permittivity ϵ and conductivity σ , hold relevant information regarding short, medium and long-term human needs. In many of these applications surface data acquisition of active source methods such as electrical resistivity (ER) and ground penetrating radar (GPR) can prove to have a

lower and more feasible deployment cost when compared with borehole methods.

ER is sensitive only to electrical conductivity while GPR is sensitive to electrical permittivity by reflectivity and velocity, and conductivity by attenuation and reflection of the excited electromagnetic wave. Full-waveform inversion (FWI) of GPR multi-offset data is an emerging technique for enhancing the resolution of electrical properties with little a-priori knowledge of the subsurface geometry with the caveat of needing an initial ray-based tomography for robust initial models (Ernst et al., 2007a,b). However, inverting with only surface acquired GPR data remains a challenge and thus limits most of the current applications in which GPR is commonly used.

Similar to seismic full-waveform inversion, two main challenges that must be resolved for GPR-FWI are the lack of low frequencies and the presence of attenuation in the data. Fortunately, ER can be used to enhance GPR because it is directly sensitive to low spatial frequencies in electrical conductivity and is directly linked to the GPR governing physics by Maxwell's equations. In this work we combine the two methods and make the assumption that electrical properties are not frequency dependent. Although this is not true in general, in Domenzain et al. (2019) we note that for a variety of relevant earth materials, the (real) effective conductivity and the DC conductivity differ by a factor of less than 5. Hence, assuming frequency independent electrical parameters serves as a starting point to test the enhancement of the spatial resolution in our inversions.

In Domenzain et al. (2019) we developed a joint inversion scheme of GPR and ER data that uses the full two dimensional physics of Maxwell's equations. The inversion accounts for the sensitivities of GPR and ER data in each iteration of an adjoint method based inversion. The GPR source wavelet is assumed known and kept constant throughout the inversion. However, it is noted that existing methods to solve for the GPR source wavelet (Pratt et al., 1998; Ernst et al., 2007a) can easily be applied

*Boise State University, Department of Geosciences, Boise, Idaho. †Colorado School of Mines, Department of Geophysics, Golden, Colorado.

to our scheme. We tested our joint inversion scheme in two synthetic examples showing enhancements when compared to individual GPR and ER inversions. The recovered conductivity was improved through joint inversion because the ER data enhanced amplitude detection and the GPR improved spatial resolution. Thus, the recovered conductivity benefits from the complementary resolution of GPR and ER data. Moreover, neither data resolution is lost.

In this work we address joint inversion of GPR and ER data when the conductivity in the subsurface is strong, i.e. for values above 10 mS/m where GPR attenuation is high and the signal-to-noise ratio in the GPR data is close to one. Unfortunately, if the attenuation is too strong the GPR data will miss reflection events that hold meaningful information of the subsurface. In this situation (Domenzain et al., 2019) we find that even though the recovered conductivity is better resolved by using both GPR and ER data, the recovered permittivity lacks the correct amplitude and misses long wavelength resolution. Fortunately, we can improve our joint inversion scheme with existing inversion methods. Specifically, we use methods that (i) enhance the low frequency content of the GPR waveform and (ii) exploit structural similarities of the subsurface parameters.

Methods developed for seismic FWI (Bozdağ et al., 2011; Liu and Zhang, 2017) can be used to enhance low frequency content in GPR-FWI. In the context of seismic FWI it is well known that low frequencies in the waveform data help the inversion avoid local minima (Virieux and Operto, 2009; Baeten et al., 2013). In Bozdağ et al. (2011) the authors propose using the analytic signal of the observed waveform in order to isolate the instantaneous phase and amplitude (i.e. envelope) information of the data and modify the FWI objective function accordingly. In Liu and Zhang (2017) the authors join first arrival travel-time with early arrival envelope data to build a rich low spatial-frequency initial velocity model that is then used in the FWI routine. Both works find that the low frequency content of the envelope waveform data is good for enhancing the low frequency spatial content of the recovered velocity. In this work we use the envelope waveform data of GPR and further join it with ER data to alleviate low spatial frequencies in both electrical permittivity and conductivity.

Inversion methods that assume structural similarities of the target subsurface parameters (Haber and Oldenburg, 1997; Gallardo and Meju, 2003) can be used to further improve our joint inversion algorithm by letting the ER data inform the GPR data in regions of high attenuation. Assuming structural similarities in target subsurface parameters allows different geophysical data with varying spatial and physical sensitivities inform each other where to look for a solution that more accurately resembles reality if the structural similarity holds true. In Gallardo and Meju (2003) the authors choose the cross-gradient operator as a structural constraint and successfully apply it

to real seismic and ER data. In this work we show that by assuming structural similarities between electrical permittivity and conductivity we can use the cross-gradient operator for filling in amplitude and spatial-frequency content to our solutions while still using forward and inverse models that take into account the full physics of Maxwell's equations.

Since then different types of geophysical data have been used in this context (Gallardo and Meju, 2007; Fregoso and Gallardo, 2009; Gross, 2019). Most relevant to our study are the works of Linde et al. (2006) and Doetsch et al. (2010) which use borehole GPR and ER data to solve for electrical permittivity and conductivity. All of these works rely on a linearization of one or both forward models and clear access to the sensitivity matrices of the data, which in case of time-domain FWI the latter is computationally expensive. In Hu et al. (2009) the authors combine seismic and controlled-source electromagnetic data to solve for compressional velocity and electrical conductivity in a Gauss-Newton inversion while enforcing the cross-gradient constraint. They employ adjoint based methods for computing the sensitivity matrices of the data with the computational burden of storing and inverting the Hessian of the objective functions. In this work we compute the gradients of the objective functions using adjoint based methods and relieve the need to store and compute the Hessian of the objective functions.

We begin with a brief review of the two-dimensional physics of the forward models for GPR and ER, and objective functions for the GPR and ER inversions. Then we review our joint inversion scheme from Domenzain et al. (2019) and define three new joint inversion schemes designed to manage attenuation and enhance low frequencies. Finally, we test our joint inversions on three synthetic subsurface models designed to challenge the spatial and amplitude resolution of GPR and ER sensitivities. The first two models illustrate our method with a centered box anomaly where the improvements of our methods are clear. The third model is based on an alluvial aquifer located at the Boise Hydrological Research Site (Barrash and Clemo, 2002). This model contains realistic electrical parameters and a subsurface geometry that loosely resembles previous GPR multi-offset data imaging done on this site (Bradford et al., 2009b). Testing the attributes and limitations of our method on synthetic data is crucial for assessing the viability of our results when used on field data.

GPR AND ER FORWARD MODELS AND INVERSIONS

We briefly recall the governing equations, forward models and objective functions for the GPR and ER experiments. Our physical models assume isotropic physical properties and a 2D subsurface geometry where the parameters are constant along the y -axis. These assumptions are made for ease of computations of our forward models and not crucial for our inversion schemes. Both the GPR and ER

forward models are discretized on the same computational grid. Gradients of the objective function with respect to the parameters are given in Appendix A and a full discussion is found in Domenzain et al. (2019).

Ground penetrating radar

The two-dimensional physics of the GPR experiment are given by,

$$\begin{pmatrix} \mu_o & 0 & 0 \\ 0 & \mu_o & 0 \\ 0 & 0 & \varepsilon \end{pmatrix} \begin{pmatrix} \dot{H}_z \\ -\dot{H}_x \\ \dot{u} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \partial_x \\ 0 & 0 & \partial_z \\ \partial_x & \partial_z & 0 \end{pmatrix} \begin{pmatrix} H_z \\ -H_x \\ u \end{pmatrix} - \sigma \begin{pmatrix} 0 \\ 0 \\ u \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ -J_y \end{pmatrix}, \quad (1)$$

where u is the electric field component in the y direction, (H_x, H_z) are the magnetic field components in the x and z direction, J_y is the source term, ε is the electric permittivity, σ is the electric conductivity and μ_o is the magnetic permeability which we assume constant and equal to the permeability of free space. Let ε_o denote the electrical permittivity of free space. From now on, we will refer to the relative permittivity $\varepsilon_r = \varepsilon/\varepsilon_o$ simply as permittivity. We discretize equation 1 by

$$\begin{aligned} \mathbf{u} &= \mathbf{L}_w \mathbf{s}_w, \\ \mathbf{d}_w^s &= \mathbf{M}_w \mathbf{u} \end{aligned} \quad (2)$$

where \mathbf{L}_w is the discretized differential (time marching) operator of equation 1, \mathbf{u} is the electric field y component defined in space and time, \mathbf{s}_w is the source term, \mathbf{M}_w is the measuring operator, and $\mathbf{d}_w^s = \mathbf{M}_w \mathbf{u}$ is the data of the experiment, i.e. a common-source gather. The discretized solution of equation 2 is described in detail in Domenzain et al. (2019).

We make note that from this point forward we will refer to operators and variables in capital and lower case letters respectively, and distinguish continuous and discrete mathematics in normal and bold font respectively. A complete table of relevant notation can be found in Table 1.

Electrical resistivity

The two-dimensional physics of the ER experiment are given by the steady state Maxwell's equations where Ohm's law holds (Pidlisecky et al., 2007),

$$-\nabla \cdot \sigma \nabla \varphi = \mathbf{i}(\delta(x - s_+) - \delta(x - s_-)), \quad (3)$$

where φ is the electric potential, \mathbf{i} is the current intensity and s_{\pm} is the source-sink position. We write the discretized version of equation 3 as,

$$\begin{aligned} \mathbf{L}_{dc} \varphi &= \mathbf{s}_{dc}, \\ \mathbf{d}_{dc}^s &= \mathbf{M}_{dc} \varphi, \end{aligned} \quad (4)$$

Symbol	Meaning	Note
ε_r	Discretized relative permittivity	
σ	Discretized conductivity	
\mathbf{L}	Discretized differential operator	
\mathbf{s}	Discretized source	
\mathbf{M}	Discretized measuring operator	
\mathbf{d}	Synthetic data	
\mathbf{e}	Residual of synthetic vs observed data	
Θ	Objective function	
\mathbf{v}	Discretized adjoint field	
\mathbf{g}	Gradient of objective function	
α	Step size for \mathbf{g}	
\mathbf{u}	Electric wavefield on the y component	
$\dot{\mathbf{u}}$	finite-difference time derivative of \mathbf{u}	
$\Delta\sigma_w$	GPR conductivity update	Used for GPR and ER
$\Delta\varepsilon_r$	GPR permittivity update	
β	GPR envelope weight	
φ	Electric potential	
\mathbf{S}_{dc}	The matrix $-((\nabla_\sigma \mathbf{L}_{dc}) \varphi)^\top$	Only ER
$\Delta\sigma_{dc}$	ER conductivity update	
$\Delta\sigma$	Joint conductivity update	
a_w, a_{dc}	Weights to regulate $\Delta\sigma_w$ and $\Delta\sigma_{dc}$	
c	Step size for $\Delta\sigma$	
$\Delta\sigma_{\tau,o}$	Cross-gradient conductivity update	
$\Delta\varepsilon_{r,\tau,o}$	Cross-gradient permittivity update	
$b_{\varepsilon_r}, b_\sigma$	Weights to regulate $\Delta\varepsilon_{r,\tau,o}$ and $\Delta\sigma_{\tau,o}$	Used for the joint update

Table 1: Reference for the notation used in the discretized inverse problems. Symbols common in both GPR and ER experiments are stripped from their subscripts to avoid clutter.

where \mathbf{L}_{dc} is the discretized differential operator of equation 3, $\boldsymbol{\varphi}$ is the electric potential, \mathbf{s}_{dc} is the source term, \mathbf{M}_{dc} is the measuring operator that computes observed voltages, and \mathbf{d}_{dc}^s is the data of the experiment for one source. The discretized solution of equation 4 is described in detail in Domenzain et al. (2019).

GPR inversion

The GPR inversion algorithm finds parameters $\boldsymbol{\varepsilon}_{r*}$ and $\boldsymbol{\sigma}_*$ that satisfy,

$$\{\boldsymbol{\varepsilon}_{r*}, \boldsymbol{\sigma}_*\} = \arg \min \Theta_w(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}; \mathbf{d}_w^o), \quad (5)$$

where the subscript * denotes the imaged parameters and \mathbf{d}_w^o denotes all the observed GPR data. From now on we denote the electrical permittivity and conductivity in bold font to emphasize these parameters are discretized and in matrix form. We have,

$$\Theta_w = \frac{1}{n_s} \sum_s \Theta_w^s, \quad (6)$$

where s indexes the sources, n_s denotes the total number of sources, and

$$\Theta_w^s = \frac{\|\mathbf{e}_w\|_2^2}{\|\mathbf{d}_w^{o,s}\|_2^2}, \quad (7)$$

where $\mathbf{d}_w^{o,s}$ is the observed data for one source and $\mathbf{e}_w = \mathbf{d}_w^s - \mathbf{d}_w^{o,s}$ is the residual of the modeled and observed data. The details for computing the gradient of Θ_w with respect to $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ can be found in Appendix A.

ER inversion

The ER inversion algorithm finds $\boldsymbol{\sigma}_*$ that satisfies,

$$\boldsymbol{\sigma}_* = \arg \min \Theta_{dc}(\boldsymbol{\sigma}; \mathbf{d}_{dc}^o), \quad (8)$$

where \mathbf{d}_{dc}^o is all of the ER data. We have,

$$\Theta_{dc} = \frac{1}{n_s} \sum_s \Theta_{dc}^s, \quad (9)$$

where s indexes the source, n_s denotes the total number of sources, and

$$\Theta_{dc}^s = \frac{\|\mathbf{e}_{dc}\|_2^2}{\|\mathbf{d}_{dc}^{o,s}\|_2^2}. \quad (10)$$

We denote $\mathbf{d}_{dc}^{o,s}$ the observed data for one source and $\mathbf{e}_{dc} = \mathbf{d}_{dc}^s - \mathbf{d}_{dc}^{o,s}$ the residual of the modeled and observed data. The details for computing the gradient of Θ_{dc} with respect to $\boldsymbol{\sigma}$ can be found in Appendix A.

JOINT INVERSIONS

Joint inversion of ER and GPR data

The objective function for our joint inversion is,

$$\{\boldsymbol{\varepsilon}_{r*}, \boldsymbol{\sigma}_*\} = \arg \min \Theta_w(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}; \mathbf{d}_w^o) + \Theta_{dc}(\boldsymbol{\sigma}; \mathbf{d}_{dc}^o). \quad (11)$$

We optimize 11 using gradient descent by first computing the descent directions for $\boldsymbol{\sigma}$: $\Delta\boldsymbol{\sigma}_w$ and $\Delta\boldsymbol{\sigma}_{dc}$ for both Θ_w and Θ_{dc} respectively, and then take a weighted average of these descent directions to update $\boldsymbol{\sigma}$; we then compute the descent direction $\Delta\boldsymbol{\varepsilon}_r$ and update $\boldsymbol{\varepsilon}_r$. Figure 1 shows a code-flow diagram of this process. We follow Domenzain et al. (2019) and briefly explain how these updates and joining-weights are computed.

After all the gradients for all sources are computed the update directions are,

$$\Delta\boldsymbol{\sigma}_w = -\frac{1}{n_w} \sum_{s=1}^{n_w} \alpha_\sigma^s \mathbf{g}_{w,\sigma}^s, \quad (12)$$

$$\Delta\boldsymbol{\sigma}_{dc} = -\frac{1}{n_{dc}} \sum_{s=1}^{n_{dc}} \alpha_{dc}^s \mathbf{g}_{dc}^s, \quad (13)$$

$$\Delta\boldsymbol{\varepsilon}_r = -\frac{1}{n_w} \sum_{s=1}^{n_w} \alpha_{\varepsilon_r}^s \mathbf{g}_{\varepsilon_r}^s, \quad (14)$$

where α_σ^s , α_{dc}^s , and $\alpha_{\varepsilon_r}^s$ are computed as in Domenzain et al. (2019). After $\Delta\boldsymbol{\sigma}_w$ and $\Delta\boldsymbol{\sigma}_{dc}$ have been computed they are joined by weights a_w and a_{dc} ,

$$\Delta\boldsymbol{\sigma} = a_w \Delta\boldsymbol{\sigma}_w + a_{dc} \Delta\boldsymbol{\sigma}_{dc}, \quad (15)$$

we then normalize $\Delta\boldsymbol{\sigma}$ by its largest amplitude and finally write,

$$\Delta\boldsymbol{\sigma} \leftarrow c \Delta\boldsymbol{\sigma}, \quad (16)$$

where c is the geometric mean of the maximum amplitudes of $\Delta\boldsymbol{\sigma}_w$ and $\Delta\boldsymbol{\sigma}_{dc}$ prior to normalization. The driving purpose of the weights a_w and a_{dc} is of letting both updates $\Delta\boldsymbol{\sigma}_w$ and $\Delta\boldsymbol{\sigma}_{dc}$ always contribute to $\Delta\boldsymbol{\sigma}$ in proportion to their objective function value at a given iteration. Figure 2 shows the shape as a function of iterations of the weights a_w and a_{dc} should have: a bow-tie shape where at early iterations a_w dominates and at later iterations a_{dc} takes over. For a full discussion on the weights a_w and a_{dc} see Domenzain et al. (2019).

In order to enforce positivity constraints the parameters are updated as (Meles et al., 2010),

$$\boldsymbol{\sigma} \leftarrow \boldsymbol{\sigma} \odot \exp(\boldsymbol{\sigma} \odot \Delta\boldsymbol{\sigma}), \quad (17)$$

$$\boldsymbol{\varepsilon}_r \leftarrow \boldsymbol{\varepsilon}_r \odot \exp(\boldsymbol{\varepsilon}_r \odot \Delta\boldsymbol{\varepsilon}_r). \quad (18)$$

We will refer to this inversion method (i.e. optimizing equation 11) as *Joint*.

Joint inversion of GPR envelope and ER data

We begin with a description of GPR envelope inversion which exploits the low frequency content of the GPR data. Similar to GPR inversion we find $\boldsymbol{\varepsilon}_{r*}$ and $\boldsymbol{\sigma}_*$ but with the objective function $\tilde{\Theta}_w$ (Bozdag et al., 2011; Liu and Zhang, 2017),

$$\{\boldsymbol{\varepsilon}_{r*}, \boldsymbol{\sigma}_*\} = \arg \min \tilde{\Theta}_w(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}; \mathbf{d}_w^o, \mathbf{d}_{w,a}^o), \quad (19)$$

$$\tilde{\Theta}_w = \Theta_w(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}; \mathbf{d}_w^o) + \Theta_{w,a}(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}; \mathbf{d}_{w,a}^o),$$

where $\mathbf{d}_{w,a}^o$ is the envelope of the observed data using the Hilbert transform and,

$$\begin{aligned}\Theta_{w,a} &= \frac{1}{n_s} \sum_s \Theta_{w,a}^s, \\ \Theta_{w,a}^s &= \frac{\|\mathbf{e}_{w,a}\|_2^2}{\|\mathbf{d}_{w,a}^o\|_2^2},\end{aligned}\quad (20)$$

where s indexes the sources. We optimize $\tilde{\Theta}_w$ using gradient descent and regulate how much information $\Theta_{w,a}$ contributes to the inversion by weighing the gradients of Θ_w and $\Theta_{w,a}$ differently. The gradients of $\Theta_{w,a}$ with respect to the parameters $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ are computed using a full-waveform approach where a different adjoint source has to be used for $\Theta_{w,a}$ as explained in Bozdağ et al. (2011) and reproduced in Appendix B for completeness.

For the sake of clarity we illustrate the optimization procedure for just $\boldsymbol{\varepsilon}_r$. For one source, let $\tilde{\mathbf{g}}_{\varepsilon_r}^s$, $\mathbf{g}_{\varepsilon_r}^s$ and $\mathbf{g}_{\varepsilon_r,a}^s$ be the gradients of $\tilde{\Theta}_w^s$, Θ_w^s and $\Theta_{w,a}^s$, where the last two are computed as in equations A-3 and B-10 respectively. We have,

$$\tilde{\mathbf{g}}_{\varepsilon_r}^s = \mathbf{g}_{\varepsilon_r}^s + \beta_{\varepsilon_r} \mathbf{g}_{\varepsilon_r,a}^s, \quad (21)$$

where the gradients $\mathbf{g}_{\varepsilon_r}^s$ and $\mathbf{g}_{\varepsilon_r,a}^s$ are assumed normalized in amplitude and β_{ε_r} is a fixed scalar quantity for all sources and all iterations. The weight β_{ε_r} regulates how much we boost the low frequency content of the observed GPR data. Our numerical results show that a larger value of β_{ε_r} gives better depth resolution with the caveat of losing spatial resolution. However if the value of β_{ε_r} is too large the inversion might strongly favor the low spatial-frequency content over the high spatial-frequency content, thus not giving accurate results.

Once $\tilde{\mathbf{g}}_{\varepsilon_r}^s$ has been computed we find the step-size $\alpha_{\varepsilon_r}^s$ as detailed in Domenzain et al. (2019). After $\tilde{\mathbf{g}}_{\varepsilon_r}^s$ and $\alpha_{\varepsilon_r}^s$ have been computed for all sources the permittivity update is,

$$\Delta\boldsymbol{\varepsilon}_r = -\frac{1}{n_w} \sum_{s=1}^{n_w} \alpha_{\varepsilon_r}^s \tilde{\mathbf{g}}_{\varepsilon_r}^s. \quad (22)$$

Analogous to $\boldsymbol{\varepsilon}_r$, the update for $\boldsymbol{\sigma}$ is,

$$\tilde{\mathbf{g}}_{\boldsymbol{\sigma}} = \mathbf{g}_{w,\boldsymbol{\sigma}}^s + \beta_{\boldsymbol{\sigma}} \mathbf{g}_{\boldsymbol{\sigma},a}^s, \quad (23)$$

$$\Delta\boldsymbol{\sigma}_w = -\frac{1}{n_w} \sum_{s=1}^{n_w} \alpha_{\boldsymbol{\sigma}} \tilde{\mathbf{g}}_{\boldsymbol{\sigma}}, \quad (24)$$

where $\beta_{\boldsymbol{\sigma}}$ is a fixed scalar quantity, $\mathbf{g}_{w,\boldsymbol{\sigma}}^s$ and $\mathbf{g}_{\boldsymbol{\sigma},a}^s$ are computed as in equations A-2 and B-9 respectively, and $\tilde{\mathbf{g}}_{\boldsymbol{\sigma}}$ is assumed normalized in amplitude. Similarly to β_{ε_r} a larger value of $\beta_{\boldsymbol{\sigma}}$ will result in better depth resolution.

The weights β_{ε_r} and $\beta_{\boldsymbol{\sigma}}$ play an important role in recovering the subsurface parameters. In our numerical results we have found that when the GPR data has a small signal-to-noise ratio it is beneficial to use values close to one and when the signal-to-noise ratio is large, values smaller than one give better results. However, regardless on how good the signal-to-noise ratio is in the GPR data using the ER data in a joint inversion proves to have better results with

comparatively stronger results when the GPR data exhibits strong attenuation.

We define our joint inversion of GPR envelope and ER data by minimizing the following objective function,

$$\{\boldsymbol{\varepsilon}_{r*}, \boldsymbol{\sigma}_*\} = \arg \min \tilde{\Theta}_w(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}; \mathbf{d}_w^o, \mathbf{d}_{w,a}^o) + \Theta_{dc}(\boldsymbol{\sigma}; \mathbf{d}_{dc}^o). \quad (25)$$

At a given iteration of our joint inversion (whose workflow is as in Figure 1) we replace $\Delta\boldsymbol{\varepsilon}_r$ and $\Delta\boldsymbol{\sigma}_w$ by those computed in equations 22 and 24. The updated values for $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}_r$ are made as in equations 17 and 18.

We will refer to this inversion method (i.e. optimizing equation 25) as JEN.

Joint inversion with cross-gradients

In this section we assume electrical permittivity and conductivity share structural properties. At a given iteration we want the structure of $\boldsymbol{\varepsilon}_r$ to be shared onto $\boldsymbol{\sigma}$ and vice-versa, and we want to do so by respecting the different concavities $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ may have. For this reason we choose the discrete cross-gradient operator $\boldsymbol{\tau}$ as a measure of structure Gallardo and Meju (2003),

$$\boldsymbol{\tau}(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}) = \nabla_{\mathbf{x}} \boldsymbol{\varepsilon}_r \times \nabla_{\mathbf{x}} \boldsymbol{\sigma}, \quad (26)$$

where $\nabla_{\mathbf{x}}$ denotes the discretized finite-difference spatial operator (∂_x, ∂_z) , and minimize the objective function Θ_{τ} ,

$$\Theta_{\tau}(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}) = \frac{1}{2} \|\boldsymbol{\tau}\|_2^2. \quad (27)$$

Because we are modeling the full physics of both the GPR and ER experiments and we compute the gradients of our objective functions using an FWI and adjoint method approach, our method differs from the original method of Gallardo and Meju (2003) since we choose to not compute the sensitivity matrices of our data. The result is that at each iteration of our joint inversion (whose workflow is shown in Figure 2) we optimize equation 27 using a Gauss-Newton approach from which we only use the *master* updates $\Delta\boldsymbol{\sigma}_{\tau,o}$ and $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$. These updates are the cumulative sum of all updates done in the Gauss-Newton optimization routine. For example, let $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$ and $\Delta\boldsymbol{\varepsilon}_{r,\tau}$ be the Gauss-Newton updates of the first and second iteration respectively for optimizing Θ_{τ} with respect to $\boldsymbol{\varepsilon}_r$. Before the third iteration, the master update $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$ takes the form,

$$\Delta\boldsymbol{\varepsilon}_{r,\tau,o} \leftarrow \Delta\boldsymbol{\varepsilon}_{r,\tau,o} + \Delta\boldsymbol{\varepsilon}_{r,\tau}. \quad (28)$$

This procedure is then repeated at each iteration. The full details of optimizing equation 27 and computing $\Delta\boldsymbol{\sigma}_{\tau,o}$ and $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$ are explained in Appendix C.

We observe that minimizing Θ_{τ} in this way (1) has good potential for a well posed problem because the number of data points is equal to the number of unknowns (all the points in our model domain), (2) is relatively cheap in computation time and memory, (3) can be done by modifying both $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ or by keeping one fixed and only

modifying the other, and (4) enables us to port the information of minimizing Θ_τ into our scheme for optimizing Θ_w and Θ_{dc} without having to use second order optimization methods, i.e. the Hessians of Θ_w and Θ_{dc} .

Figure 4 gives an example of the different possibilities for minimizing Θ_τ outlined in observation (3). Given hypothetical values for $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ in Figures 4a and 4b, at a given iteration we minimize Θ_τ in three different ways. In Figures 4c and 4d we update $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$, in Figure 4e we fix $\boldsymbol{\sigma}$ and update $\boldsymbol{\varepsilon}_r$, and in Figure 4f we fix $\boldsymbol{\varepsilon}_r$ and update $\boldsymbol{\sigma}$. Note that in this example both $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ have different concavities and different shapes, i.e. $\boldsymbol{\sigma}$ is wider than $\boldsymbol{\varepsilon}_r$, mimicking the different resolutions our joint inversion is able to obtain from these two different parameters. The dashed circles are of fixed radii in all panels and serve as markers for the underlying shapes.

When optimizing Θ_τ for both $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}_r$ as shown in Figures 4c and 4d, both $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}_r$ are modified and reshaped to look more like one another since they are jointly updated. Figure 4e shows $\boldsymbol{\varepsilon}_r$ expanding towards the outer circle, appearing even more similar to $\boldsymbol{\sigma}$ in Figure 4b than that of Figure 4c. Figure 4f shows $\boldsymbol{\sigma}$ contracting into the inner circle, appearing even more similar to $\boldsymbol{\varepsilon}_r$ in Figure 4a than that of Figure 4d.

Depending on the subsurface material properties, the sensitivities of the GPR and ER data might resolve better at earlier iterations either $\boldsymbol{\varepsilon}_r$ or $\boldsymbol{\sigma}$. Whichever subsurface parameter is best resolved first should inform the other about its structural properties. Because of this reason and observations (1)-(4) above we choose to optimize Θ_τ twice per iteration: once modifying $\boldsymbol{\sigma}$ and keeping $\boldsymbol{\varepsilon}_r$ fixed and a second time modifying $\boldsymbol{\varepsilon}_r$ and keeping $\boldsymbol{\sigma}$ fixed. Each optimization has unique weights b_σ and b_{ε_r} that identify how much confidence we give to the current solutions of either $\boldsymbol{\varepsilon}_r$ or $\boldsymbol{\sigma}$.

We define our joint GPR and ER with cross-gradient by minimizing the following objective function,

$$\{\boldsymbol{\varepsilon}_{r*}, \boldsymbol{\sigma}_*\} = \arg \min \Theta_w(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}; \mathbf{d}_w^o) + \Theta_{dc}(\boldsymbol{\sigma}; \mathbf{d}_{dc}^o) + \Theta_\tau(\boldsymbol{\varepsilon}_r, \boldsymbol{\sigma}). \quad (29)$$

At each iteration of our joint inversion we begin with estimates of $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$. The joint update for the conductivity first involves keeping $\boldsymbol{\varepsilon}_r$ fixed and computing the update $\Delta\boldsymbol{\sigma}_{\tau,o}$ given by equation C-7 that optimizes Θ_τ . Then we compute the weight b_σ and scale $\Delta\boldsymbol{\sigma}_{\tau,o}$,

$$b_\sigma = \left(h_\sigma \frac{a_{dc}}{a_w} - (h_\sigma - d_\sigma) a_{dc} \bullet \right) a_w, \quad (30)$$

$$\Delta\boldsymbol{\sigma}_{\tau,o} \leftarrow b_\sigma \Delta\boldsymbol{\sigma}_{\tau,o},$$

where $a_{dc} \bullet$ is the value of a_{dc} in the first iteration, $\Delta\boldsymbol{\sigma}_{\tau,o}$ is assumed normalized in amplitude. The scalars d_σ and h_σ control how early and how much in the joint inversion should the structural information of $\boldsymbol{\varepsilon}_r$ is to be imprinted in $\boldsymbol{\sigma}$.

Figure 2 depicts the optimal path of b_σ throughout the iterations. During early iterations b_σ is small because not enough structure has been recovered on $\boldsymbol{\varepsilon}_r$, but at late

iterations b_σ is larger because $\boldsymbol{\varepsilon}_r$ is closer to its true solution. The value of b_σ at a given iteration is a measure of how much confidence we have on the structure of the current solution for $\boldsymbol{\varepsilon}_r$: the larger b_σ the more confidence we have on $\boldsymbol{\varepsilon}_r$.

We note that the upward trend of b_σ over iterations can only be achieved if

$$h_\sigma \geq d_\sigma > 0, \quad (31)$$

which also forces b_σ to plateau to the value h_σ in late iterations so as to inhibit dominance of the structural assumption and let the physics of our inversions assume control. The purpose of d_σ is to control the value of b_σ for the first iteration: $b_\sigma = d_\sigma a_{dc} \bullet$.

The new update $\Delta\boldsymbol{\sigma}_{\tau,o}$ is now passed to the GPR and ER optimization routines before the step-sizes of the gradients are computed,

$$\begin{aligned} \mathbf{g}_{w,\sigma}^s &\leftarrow \mathbf{g}_{w,\sigma}^s + \Delta\boldsymbol{\sigma}_{\tau,o}, \\ \mathbf{g}_{dc}^s &\leftarrow \mathbf{g}_{dc}^s + \Delta\boldsymbol{\sigma}_{\tau,o}, \end{aligned} \quad (32)$$

where both $\mathbf{g}_{w,\sigma}^s$ and \mathbf{g}_{dc}^s are assumed normalized in amplitude. The step-sizes of the gradients $\mathbf{g}_{w,\sigma}^s$ and \mathbf{g}_{dc}^s are computed as described in Domenzain et al. (2019) and the updates $\Delta\boldsymbol{\sigma}_w$ and $\Delta\boldsymbol{\sigma}_{dc}$ are computed as in equations 12 and 13. Finally, the updated value for $\boldsymbol{\sigma}$ is calculated as in equation 17.

Figure 2 shows a code-flow diagram of our joint inversion with the cross-gradient. The next step in our joint inversion is the structural update to $\boldsymbol{\varepsilon}_r$ which is analogous to the update we just computed for $\boldsymbol{\sigma}$. We keep $\boldsymbol{\sigma}$ fixed, compute $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$ given by equation C-6, compute the weight b_{ε_r} and scale $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$,

$$\begin{aligned} b_{\varepsilon_r} &= \left(h_{\varepsilon_r} \frac{a_{dc}}{a_w} - (h_{\varepsilon_r} - d_{\varepsilon_r}) a_{dc} \bullet \right) a_w, \\ \Delta\boldsymbol{\varepsilon}_{r,\tau,o} &\leftarrow b_{\varepsilon_r} \Delta\boldsymbol{\varepsilon}_{r,\tau,o}, \end{aligned} \quad (33)$$

where $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$ is normalized in amplitude. The new update $\Delta\boldsymbol{\varepsilon}_{r,\tau,o}$ is now passed to the GPR optimization routine before the step-size of the gradient is computed by

$$\mathbf{g}_{\varepsilon_r}^s \leftarrow \mathbf{g}_{\varepsilon_r}^s + \Delta\boldsymbol{\varepsilon}_{r,\tau,o}, \quad (34)$$

where $\mathbf{g}_{\varepsilon_r}^s$ is assumed normalized in amplitude. The updated value for $\boldsymbol{\varepsilon}_r$ is calculated by equation 18 where the update $\Delta\boldsymbol{\varepsilon}_r$ is given in equation 14.

The weights h_{ε_r} and d_{ε_r} are not necessarily equal to h_σ and d_σ , but b_{ε_r} must follow a similar shape as b_σ (see Figure 2). Similar to b_σ , the value of b_{ε_r} at a given iteration is a measure of how much confidence we have on the structure of the current solution for $\boldsymbol{\varepsilon}_r$: the larger b_{ε_r} the more confidence we have on $\boldsymbol{\varepsilon}_r$.

Because h_σ and h_{ε_r} regulate how large b_σ and b_{ε_r} can become over the course of iterations, we propose two general rules on choosing h_σ and h_{ε_r} based on how much conductivity is present in the subsurface:

1. if conductivity is low h_{ε_r} should be small and h_σ large,

2. if conductivity is high h_σ should be small and h_{ε_r} large.

We recognize that in a real scenario we might not know a-priori the conductivity of the subsurface, however we can obtain a good enough approximation for determining h_σ and h_{ε_r} by observing the ER pseudo-sections and assessing how many reflection events are visible in the GPR shot-gathers.

We will refer to this inversion method (i.e. optimizing equation 29) as *JOIX*.

Joint inversion of GPR envelope and ER data with cross-gradient

Now that we have enhanced our joint inversion of GPR and ER data (Domenzain et al., 2019) with an envelope objective function for the GPR data and with structural similarities of subsurface electrical properties, we develop a third method that joins these two enhancements into one single inversion procedure. The joint GPR envelope and ER data with cross-gradient inversion minimizes the following objective function,

$$\{\varepsilon_{r*}, \boldsymbol{\sigma}_*\} = \arg \min \tilde{\Theta}_w(\varepsilon_r, \boldsymbol{\sigma}; \mathbf{d}_w^o, \mathbf{d}_{w,a}^o) + \Theta_{dc}(\boldsymbol{\sigma}; \mathbf{d}_{dc}^o) + \Theta_\tau(\varepsilon_r, \boldsymbol{\sigma}). \quad (35)$$

At a given iteration we first compute $\Delta\boldsymbol{\sigma}_{\tau,o}$ as in equation 30, and then add this information to the gradients $\mathbf{g}_{w,\sigma}^s$ and \mathbf{g}_{dc}^s normalized in amplitude given by equations A-2 and A-4,

$$\begin{aligned} \mathbf{g}_{w,\sigma}^s &\leftarrow \mathbf{g}_{w,\sigma}^s + \beta_\sigma \mathbf{g}_{\sigma,a}^s + \Delta\boldsymbol{\sigma}_{\tau,o}, \\ \mathbf{g}_{dc}^s &\leftarrow \mathbf{g}_{dc}^s + \Delta\boldsymbol{\sigma}_{\tau,o}. \end{aligned} \quad (36)$$

Once the gradients from all sources have been computed, we find the updates $\Delta\boldsymbol{\sigma}_w$ and $\Delta\boldsymbol{\sigma}_{dc}$ as given by equations 12 and 13. Then we can compute $\Delta\boldsymbol{\sigma}$ with equation 16 and update $\boldsymbol{\sigma}$ as in equation 17.

In order to compute $\Delta\varepsilon_r$ we first compute $\Delta\varepsilon_{r,\tau,o}$ as in equation 33 and then add this information to $\mathbf{g}_{\varepsilon_r}^s$ and $\mathbf{g}_{\varepsilon_r,a}^s$ normalized in amplitude as given by equations A-3 and B-10,

$$\mathbf{g}_{\varepsilon_r}^s \leftarrow \mathbf{g}_{\varepsilon_r}^s + \beta_{\varepsilon_r} \mathbf{g}_{\varepsilon_r,a}^s + \Delta\varepsilon_{r,\tau,o}. \quad (37)$$

Once all gradients for all sources have been computed we find $\Delta\varepsilon_r$ as given by equation 14. Finally we update ε_r as in equation 18. The code-flow diagram in Figure 2 also describes this procedure with gradients computed by equations 36 and 37.

We will refer to this inversion method (i.e. optimizing equation 35) as *JENX*.

Choice of weights

In order to join the objective functions $\tilde{\Theta}_w$, Θ_{dc} and Θ_τ we have introduced 11 weights. Our joint inversion requires 5 (equation 15), the envelope inversion requires 2

(equations 21 and 23) and the cross-gradient inversion requires 4 (equations 30 and 33). Aside from the considerations given for each inversion routine, our numerical results show that when all weights are non-zero they all influence each other. In some cases the influence the weights exert on each other can lead to a different behavior in the inversion than what was explained in the previous sections.

We observe that the conductivity solution influences the permittivity solution in a stronger way than the permittivity solution influences the conductivity solution. Moreover, because of the weak sensitivity the GPR data has on the conductivity, obtaining a good solution for the conductivity is most efficiently achieved by joining the ER data (Domenzain et al., 2019). Therefore, we assume we are already satisfied with the joint weights of equation 15 and focus on improving the permittivity and conductivity solution with weights for $\tilde{\Theta}_w$ or Θ_τ .

Let us first assume the conductivity of the subsurface is low and the GPR data holds enough information for a good solution of the permittivity. If we increase β_σ or h_σ (in equations 36 and 30) for a better depth or spatial resolution of $\boldsymbol{\sigma}$ we pay the price of degrading the spatial and amplitude resolution of ε_r .

Let us now assume the conductivity in the subsurface is high and the GPR data does not hold enough information for a good estimate of the permittivity but the ER data is enough for a good solution of the conductivity. Contrary to the above scenario, in this case it is possible to exploit the good solution of $\boldsymbol{\sigma}$ and the low frequency content of the GPR data in order to improve ε_r . Our approach consists of over-weighting the envelope of the GPR data and relying on the cross-gradients to regulate the excess of the low-frequency content. We choose negative weights b_{ε_r} and b_σ for the cross-gradient updates in order to trim off the low-frequency over-fit. The use of negative weights on Θ_τ to counteract an overfit due to $\tilde{\Theta}_w$ is a novel approach to effectively using both of cross-gradients and the envelope transform as it takes into account the sensitivities of both objective functions at each iteration. The descent direction for the ε_r solution is ensured by computing the step-size for the updates with a parabolic line search as explained in Domenzain et al. (2019).

We recognize that all 11 weights were found by trial and error. In the low-conductivity scenario we followed the qualitative guidelines explained in the previous sections and shown in Figure 2. For the high-conductivity scenario the negative weights (h_{ε_r} , d_{ε_r} , h_σ and d_σ) were chosen in order for b_{ε_r} and b_σ to smoothly decrease magnitude in absolute value as iterations progressed. This choice results in more low-frequency content trim-off at early iterations and less at later iterations.

SUBSURFACE SIMULATIONS

Recovering electrical permittivity and conductivity of the subsurface using full-waveform inversion of one-sided acquired GPR data can be challenging if low frequencies

are sparse and attenuation is high. Furthermore if the subsurface geometry has velocity and attenuation anomalies larger than a wavelength of the GPR signal the data might miss amplitude information to accurately recover said anomalies. Recovering electrical conductivity of the subsurface at depth using one-sided acquired ER data is limited by needing large offsets. Furthermore if the subsurface has electrical conductivity anomalies smaller than the receiver electrode distance, the ER data cannot spatially resolve said anomaly.

Joining GPR and ER data (whose different sensitivities compliment each other by sharing electrical conductivity) can better resolve subsurface electrical properties given that both GPR and ER data hold enough information about the subsurface. However if the subsurface is poorly conductive the ER data might have little sensitivity to changes in the conductivity when compared to the GPR data. Conversely, if the subsurface is highly conductive the ER data might have a larger sensitivity to changes in the conductivity when compared to the GPR data.

In view of these observations and in an effort to keep our analysis as simple as possible we choose to test our algorithms on two synthetically designed subsurface scenarios: one with low and one with high electrical conductivity as shown in Figures 5a and 5b and Figures 5a and 5c. Both scenarios have the same subsurface geometry: an electrical velocity and conductive box-anomaly in the center and a velocity reflector at depth. The box is $1 \times 1\text{m}$ wide: two wavelengths long but just within the limit of our chosen ER experiment spatial resolution.

Finally, we implement our algorithm with all objective functions in a realistic scenario resembling an alluvial aquifer as shown in Figures 6a and 6c. Our synthetic aquifer loosely follows the subsurface geometry of the Boise Hydrogeophysical Research Site (BHR) as imaged by Bradford et al. (2009a) and mapped by Barrash and Clemo (2002). The electrical parameters resemble those of dry gravel on the shallow layer and a variety of moist sands in the deeper layers, with wetter sands (but not saturated) to the left of the model. The dipping shallow layer is at most two wavelengths deep and just within our ER spatial resolution. The wet region acts both as a strong reflector and as attenuative media for the radar data. We note that by choosing this synthetic model, our data resembles a realistic field acquisition scenario.

In an effort to clarify our method, all inversions assume the GPR source wavelet is known. Moreover, our scheme can easily incorporate radar source estimation schemes such as Pratt et al. (1998) and Ernst et al. (2007a).

In the remainder of this section we address each of our three synthetic scenarios: (i) low conductivity, (ii) high conductivity and (iii) the synthetic alluvial aquifer. For each of our synthetic models, (a) we explain the experiment design and choice of the initial models used in our inversions, (b) show results for each of our inversion schemes, and (c) discuss our results.

Experiments

Low & high conductivity

We model 250MHz GPR antennas with a Ricker wavelet source. We apply 20 equally spaced sources on the air-ground interface with source-receiver near-offset of 0.5m (approximately one wavelength) and receiver-receiver distance a quarter of a wavelength as shown in Figure 5a. For the ER experiment we use 17 electrodes placed 1m away from each other on the air-ground interface (see Figure 5b) and perform all possible dipole-dipole and Wenner array configurations.

The synthetic GPR and ER data are then given random white noise with amplitude of 10% of their standard deviation as explained in Domenzain et al. (2019). See Figures 7 and 8 for the acquired data in both scenarios. Note that for the high-conductivity scenario the signal in the GPR data is very weak, and near where the box reflection event should be the signal-to-noise ratio is almost 1, while for the low-conductivity scenario the GPR data shows strong reflections.

All inversions have a starting homogeneous model for both permittivity and conductivity: a value of 4 for permittivity, and values of 1mS/m and 5mS/m for the low and high-conductivity scenarios respectively.

Synthetic alluvial aquifer

We use the same acquisition geometry as for the low and high conductivity experiments (see Figures 5a and 5b). Given the complicated subsurface geometry, we enhance the ER experiment with all possible Schlumberger arrays. All our data are given random white noise analogous to the low and high conductivity scenarios. To aid our analysis we place boreholes B1, B2 and B3 as shown in Figures 6a and 6c.

Figure 6b shows the initial permittivity and Figure 6d the initial conductivity used in our inversions. Our numerical experiments suggest a very strong sensitivity to the first layer in our initial models throughout our inversions. We choose a smooth initial model that accurately resolves the first air-wave refraction in the GPR data and qualitatively follows the shape of the low-velocity region in length. Figure 9a gives the residual of the initial and observed GPR data: all reflection events below the first air-wave refraction are present.

In Appendix D we give the details for choosing and building our initial models. The strategy consists in perturbing the true model in two different ways. First, we smooth it enough to loose depth resolution of the first layer and lateral resolution of the low-velocity region. As a second approach, we smooth the true model below the first layer but retain the true model for the first layer. The smoothing is done with a gaussian low-pass filter in the space frequency domain with a half-width of 0.8 1/m. The initial model in Figures 6b and 6d is an intermediate step between the first and second perturbations. It

is described in Appendix D. We note that although demanding, these initial models are representations of the long wavelength structure that could realistically be obtained from reflection tomography and careful analysis of direct arrivals.

Results

Low conductivity

Figures 10a and 11a show the recovered parameters for the low-conductivity case using joint inversion of GPR and ER data and using the weights in the first column of Table 2. We see the shape and amplitude of the box recovered in the permittivity solution together with high spatial-frequency artifacts around the box mainly due to one-sided acquisition and noise in the data. The recovered conductivity also exhibits high spatial-frequency artifacts around the box and a strong amplitude from the permittivity bottom reflector due to the GPR data being unable to distinguish permittivity from conductivity reflections.

Figures 10b and 11b show the recovered parameters for the low-conductivity case using joint inversion of GPR envelope and ER data and using the weights in the second column of Table 2. In the recovered permittivity we note less high spatial-frequency artifacts than in the joint inversion case (see Figure 10a), although the price to pay is a lower resolution of the box. The recovered conductivity shows better amplitude resolution although the bottom permittivity reflector is now thicker than in the joint inversion case (see Figure 11a) due to the larger weighing of the GPR low frequency.

Figures 10c and 11c show the recovered parameters for the low-conductivity case using joint inversion of GPR and ER data with cross-gradients and using the weights in the third column of Table 2. We see the permittivity solution is very similar to the joint inversion result (Figure 10a). However, the recovered conductivity has a more even spread in amplitude resolution compared to the joint and envelope inversion and the artifact amplitude of the permittivity reflector is now less as compared with Figures 11a and b.

Figures 10d and 11d show the recovered parameters for the low-conductivity case using joint inversion of GPR envelope and ER data with cross-gradients and using the weights in the fourth column of Table 2. The permittivity solution is again very similar to the results of Figures 10a and c but the conductivity solution is now slightly better than the rest of the inversion results by having a more localized resolution around the box.

High conductivity

Figures 12a and 13a show the recovered parameters for the high-conductivity case using joint inversion of GPR and ER data with weights as in the first column of Table 3. We note very weak amplitude and low spatial-frequency resolution on the recovered permittivity due to strong attenuation and a signal-to-noise ratio equal almost to 1 in the region of the box reflection event. The recovered conductivity exhibits better low spatial-frequency content than the low-conductivity case, however, there are stronger amplitudes near the top of the box than at depth.

Figures 12b and 13b show the recovered parameters for the high-conductivity case using joint inversion of GPR envelope and ER data with weights as in the second column of Table 3. The recovered permittivity now exhibits less high-spatial frequency content than in the joint inversion of GPR and ER case (see Figure 12a) and a small increase in amplitude resolution near the box anomaly. For the recovered conductivity we note a slight increase in amplitude resolution at depth.

Figures 12c and 13c show the recovered parameters for the high-conductivity case using joint inversion of GPR and ER data with cross-gradients and weights as in the third column of Table 3. We see the improved amplitude resolution in the region where the permittivity box lies, although the overall shape is missing low spatial-frequency information. The recovered conductivity now has a better depth amplitude resolution as compared with the joint GPR and ER and joint GPR envelope and ER inversions (see Figures 13a and 13b).

Figures 12d and 13d show the recovered parameters for the high-conductivity case using joint inversion of GPR envelope and ER data with cross gradients and weights as in the fourth column of Table 3. The permittivity anomaly is now recovered with an accurate amplitude and overall correct shape, however we observe an overshoot of low spatial-frequency content as a remanent artifact from the conductivity solution and the smoothing factor in the gradients. The recovered conductivity however, is now more accurate at depth and a better overall spatial resolution than the rest of the inversions.

Synthetic alluvial aquifer

In Lavoué et al. (2014) the authors invert GPR surface acquired data of a synthetic realistic subsurface scenario. The authors use a full-waveform approach and they note that regularization is needed for constraining the conductivity solution. In this work, we apply no additional regularization of the inversion beyond the joint objective function itself and the cross-gradients constraint.

Similar to our discussion for low and high conductivity, we performed all our inversions (Joint, JEN, JOIX, JENX) on the synthetic alluvial aquifer with inversion parameters as in Table 4. Figure 14 shows the recovered permittivity and Figure 15 shows the recovered conductivity for all inversions.

In Figure 14 for all inversions we see artifact ripples in the first layer. These ripples are due to the small discrepancy between values of the true and initial model (approximately 2.5% in the first layer). Similar lower space-frequency artifacts are also present in the recovered conductivity (see Figure 15).

Throughout Figure 14 we see the effect of having such

a high impedance contrast between the first layer and the low-velocity region: one-sided acquisition struggles to resolve the immediate section of the region below the first layer. As seen in Appendix D, this effect can be drastically reduced in the entire domain if the first layer of our model is more accurately resolved in the initial models.

If the subsurface anomalies are larger than a wavelength, resolving the region of the intrusion below the

first layer can be very challenging to resolve using only GPR data. Relying on the envelope of the GPR data (Figures 14b and 14d) to correct it can cause overshooting the solution. However, by using the ER sensitivity of the conductivity and the cross-gradients constraint we help mitigate this effect. By doing so, we retain the right values of permittivity and resolve the corner of the low-velocity region, see Figure 14c and Figure 15c. The cross-gradients constraint also helps stabilize the inversion by enabling to run more iterations without strong artifacts appearing in the recovered parameters.

We show the borehole data for the JOIX inversion (see Figure 14c and Figure 15c) in Figure 16 and Figure 17 for permittivity and conductivity respectively. In Figure 16 we see that despite underestimating the parameters in the initial model, the permittivity solution accurately approximates the correct values. We also note that permittivity values at depth lack precision. However, the inversion accurately locates the location of boundaries, and it does so approximating the right impedance value.

In Figure 17 we also note a lack of accuracy at depth for the recovered conductivity. Similar to the inherent lack of sensitivity in the GPR data due to two-way travel, the ER data is mostly sensitive in an upside-down trapezoid region below the survey line. The sensitivity of the ER data is mostly appreciated in Figure 15, where the conductivity is mostly resolved in a trapezoid region. Figure 17c also exhibits the lack of GPR and ER sensitivity at depth, where although the data is sensitive to impedance contrasts, it is not capable of resolving the correct magnitude for the conductivity.

Figure 18b gives the recovered GPR data for shot-gather #7 and Figure 18c gives both the observed and recovered ER data. We note that most of the reflection events of the observed GPR data below the air-wave refraction are recovered in Figure 18b. Figure 9b shows the residual of the recovered and observed GPR data. We see that the first and second air-wave refraction are recovered, and the corner of the low-velocity region is resolved up to the noise level. At early times we also note in Figure 9b the artifact ripples in the first layer that the inversion has introduced.

Discussion

Low & high conductivity

Our numerical results show that all the different objective functions $\tilde{\Theta}_w$, Θ_{dc} and Θ_τ influence each other when compared to their individual inversions. For both the low and high-conductivity scenarios we find the best results when

combining all the objective functions noting improvements in high and low spatial-frequencies, and enhancing amplitude resolution both of the box anomaly and at depth.

In all cases we find the conductivity solutions are significantly of lower spatial resolution when compared to the permittivity solutions. This is due to the inherent spatial resolution limitations of the ER data and the attenuation driven sensitivity of the GPR data to conductivity.

In the low-conductivity scenario we observe a gradual improvement in the conductivity solution by introducing the objective functions $\tilde{\Theta}_w$, Θ_{dc} and Θ_τ . We quantify this improvement by computing the absolute RMS error of the true and recovered conductivity for each method in a region around the box-anomaly and shown in the second column of Table 5. However, the improvement in the conductivity solution slightly degrades the best result for the permittivity solution as shown in the first column of Table 5. The average of both the permittivity and conductivity RMS absolute errors is displayed in the third column of Table 5, indicating that the Joint inversion of GPR and ER data with cross-gradients gives the best overall result.

In the high-conductivity case it is more clear how both the permittivity and conductivity solutions improve when introducing all objective functions. We quantify our inversion results in Table 6, which is analogous to Table 5 but for the high-conductivity scenario. The smallest RMS errors for both parameters are given by introducing all $\tilde{\Theta}_w$, Θ_{dc} and Θ_τ objective functions.

We conclude that in the low-conductivity scenario where the GPR data is strongly sensitive to permittivity, improving the conductivity solution costs a slight degradation of the permittivity solution. In the high-conductivity scenario where the GPR data is strongly affected by attenuation (and thus a lower signal-to-noise ratio), we can improve the permittivity solution by directly using data that is not directly sensitive to permittivity, i.e. ER data using cross-gradients.

Because on average for both low and high-conductivity scenarios the best recovered parameters are obtained using all objective functions (see third column of Tables 5 and 6), given field GPR and ER data we recommend using all objective functions. In the case where the GPR data is strongly sensitive to permittivity we advise caution with overweighing the envelope gradients of $\tilde{\Theta}_w$ while more leeway can be given to Θ_τ in order to improve the conductivity solution. In the case the GPR data is weakly sensitive to permittivity, we recommend strong weighing on Θ_τ in order to exploit the ER data for the benefit of the permittivity solution.

Synthetic alluvial aquifer

Compared to the low and high conductivity examples, the initial model we used for the synthetic alluvial aquifer holds much more low-spatial frequency content of the subsurface. This mostly impacts two aspects of the inversion: 1) the initial conductivity model already describes the ER data pretty well, yielding the sensitivity of the ER is weak.

2) Using the envelope of the GPR data inhibits the FWI gradient to fully exploit high spatial-frequency features. In this case, the permittivity sensitivity given by the GPR data can be exploited to improve the spatial resolution of the recovered conductivity with the cross-gradients constraint. We find the better results by completely muting the envelope weighting. This weighting strategy is in accordance with the low and high-conductivity discussion above. The cross-gradients constraint on the permittivity enhances low spatial-frequency content on the GPR sensitivity, keeping the inversion artifact-free for more iterations.

Figure 19a shows the weights a_w and a_{dc} as a function of iterations. We choose a very small starting value for a_{dc} in order to let the GPR sensitivity resolve the missing high-spatial frequency content. In Figure 19b we see that most of the model is resolved in the first 50 iterations. The next 50 iterations resolve mostly the ER data. After 150 iterations the parameters are resolved within the resolution of our methods since no relevant change occurs. Later iterations keep improving the permittivity and conductivity solutions by filling high spatial-frequency details like for example, the corner of the low-velocity region.

CONCLUSIONS

We have developed a joint inversion algorithm for one-sided acquired full-waveform GPR and ER data. The algorithm directly joins GPR and ER data, the envelope of the GPR data, and structural information of the parameters using a modified cross-gradients approach. Our three-for-one algorithm manages how much information from each sensitivity is used in the inversion. This algorithm manages effects of strong attenuation and enhances low spatial-frequency content in the recovered electrical permittivity and conductivity.

We tested our inversion scheme on synthetic noisy data and found that even in regions of high attenuation where the GPR data has a signal-to-noise ratio close to one we are able to recover accurate enough subsurface electrical properties. In regions where the attenuation is present but not strong we are able to improve the low spatial-frequency content and accurately resolve sharp boundaries of the recovered parameters.

By joining GPR with ER data we exploit the linkage given by Maxwell's equations of electrical conductivity in both GPR and ER experiments. Borrowing from seismic FWI we use the envelope of the GPR data to better resolve amplitudes at depth and improve the low-spatial frequency content. We have modified the original cross-gradient scheme to fit with our full-physics inversion without the need for computing sensitivity matrices of the data or Hessians of the objective functions.

We note that with field data scenarios it might be the case that the more attenuation in the GPR data the more sensitive to the subsurface the ER data might be (high-conductivity scenario), and the less attenuation in the GPR data the less sensitive to the subsurface the ER data

might be (low-conductivity scenario). However, our algorithm accounts for both scenarios.

We tested our algorithm on a realistic scenario based on an alluvial aquifer deposit. We find that the choice for an initial model greatly impacts the recovered parameters. The best results were found using a smooth velocity model accurate in shallow depths. We note that although demanding, our initial models may be possible to realize with field data using existing workflows such as reflection-traveltime and ER tomography. Our regularization strategy relies on letting the GPR and ER data regularize each other, together with cross-gradients constraints on both permittivity and conductivity. Albeit an initial model, no further a priori information is needed.

Even though we have presented 2D results our algorithm can take into account 3D structure by using 3D GPR and ER forward models. An important caveat of our scheme is assuming ER and GPR are sensitive to a unique electrical conductivity, and in doing so we do not account for frequency dependent conductivity. While in some limited types of materials this approximation is reasonable, in general it is not adequate. Future work will be focused toward accounting for apparent conductivity differences at DC and radar frequencies.

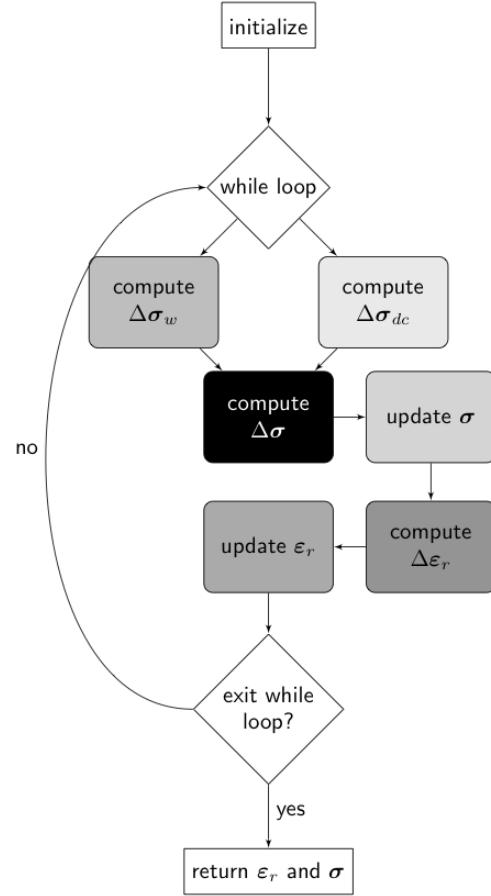


Figure 1: Inversion algorithm for Joint and JEN. We differentiate Joint and JEN by how we compute $\Delta\sigma_w$ and $\Delta\epsilon_r$.

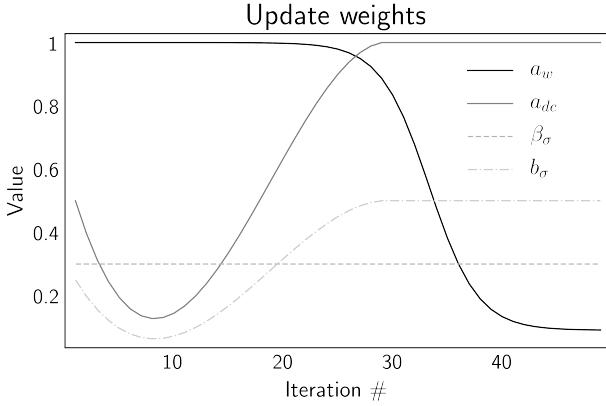


Figure 2: Qualitative optimal shape for weights throughout iterations for all inversion schemes (Joint, JEN, JOIX and JENX). Because the ER data struggles to resolve the conductivity at depth in early iterations and the GPR data first resolves the structure of the model, the weight a_w is given a larger value than a_{dc} at early iterations. Once the GPR data has resolved enough structure, the roles of a_w and a_{dc} are reversed. The envelope weights β_{ε_r} and β_σ remain constant through the inversion. The cross-gradient weights b_{ε_r} and b_σ increase their contribution through the inversion as the parameters are better resolved.

Low σ	Joint	JEN	JOIX	JENX
$a_{dc} \bullet$	0.85	0.85	0.85	0.85
\dot{a}_{dc}	3	3	3	3
$\dot{\Theta}_{dc}$	2	2	2	2
\dot{a}_w	4	4	4	4
$\dot{\Theta}_w$	0.9	0.9	0.9	0.9
β_{ε_r}		0.25		
β_σ		0.25		
h_{ε_r}			0.01	
d_{ε_r}			0.1	
h_σ				1e-3
d_σ				

Table 2: Inversion parameters for the low-conductivity scenario.

APPENDIX A

GPR AND ER GRADIENTS

We obtain the gradients $\mathbf{g}_{w,\sigma}^s$ and $\mathbf{g}_{\varepsilon_r}^s$ of Θ_w^s with respect to σ and ε_r following Meles et al. (2010) and Domenzain

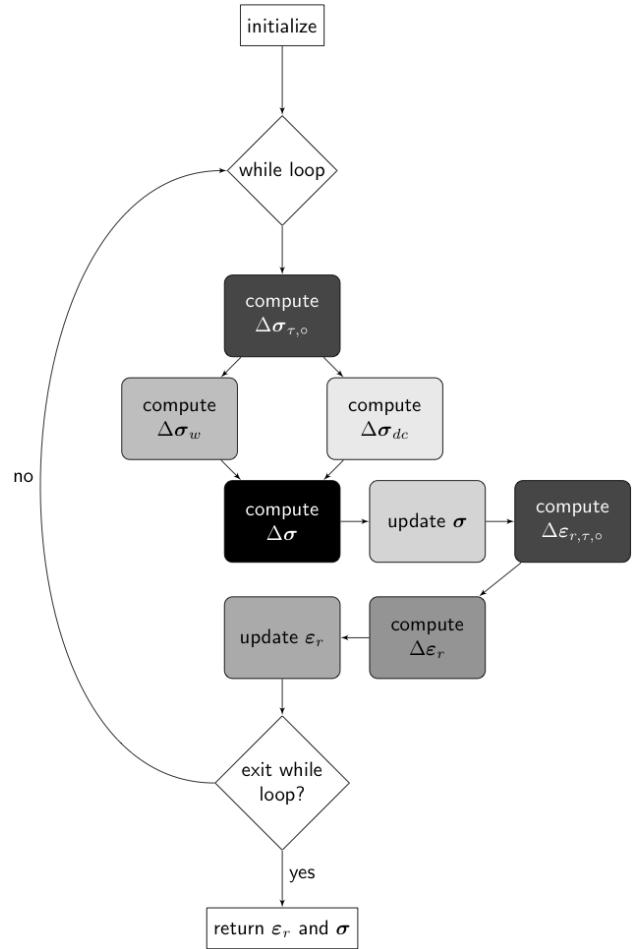


Figure 3: Inversion algorithm for JOIX and JENX. We differentiate JOIX and JENX by how we compute $\Delta\sigma_w$ and $\Delta\epsilon_r$.

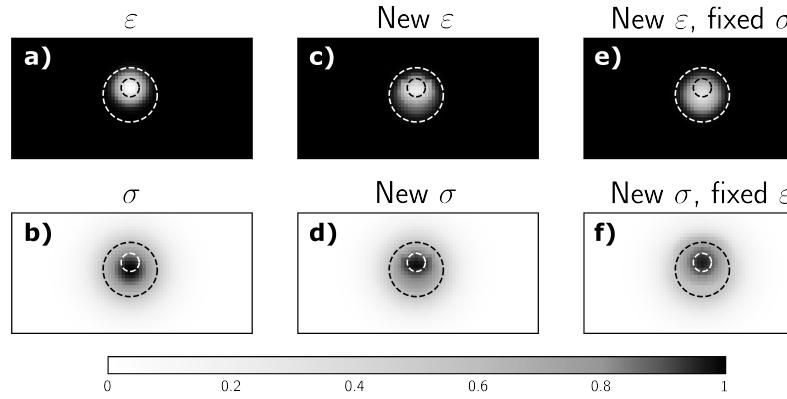


Figure 4: Illustration of cross-gradient possibilities. Given estimates ε_r and σ in **a** and **b**, Θ_τ is minimized by updating both ε_r and σ in **c** and **d**, updating ε_r and keeping σ fixed in **e**, and updating σ keeping ε_r fixed in **f**. The dashed circles are constant markers for the widths and centers of the gaussian shapes in the given estimates of ε_r and σ .

High σ	Joint	JEN	JOIX	JENX
$a_{dc} \bullet$	0.85	0.85	0.85	0.87
\dot{a}_{dc}	1.5	1.5	1.5	1.5
$\dot{\Theta}_{dc}$	1.5	1.5	1.5	1.5
\dot{a}_w	2.5	2.5	2.5	2.5
$\dot{\Theta}_w$	0.9	0.9	0.9	0.9
β_{ε_r}		1		0.5
β_σ		1		0.5
h_{ε_r}			0.2	-0.3
d_{ε_r}			0.6	-3
h_σ				-0.16
d_σ				-0.6

Table 3: Inversion parameters for the high-conductivity scenario.

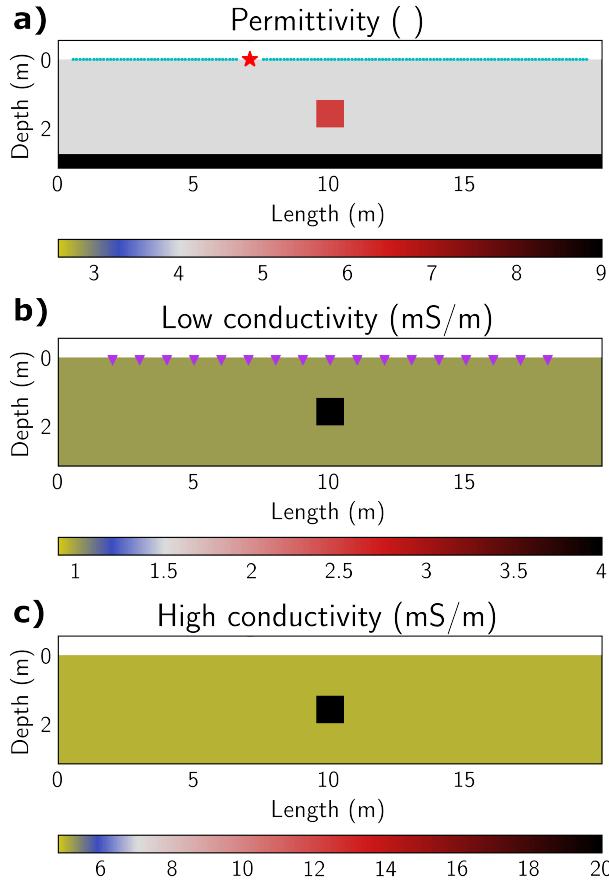


Figure 5: True permittivity **a** and conductivity for the low **b** and high **c** conductivity scenario. In **a**, an example of GPR receivers (cyan) and source (red). ER electrodes are shown purple in **b**.

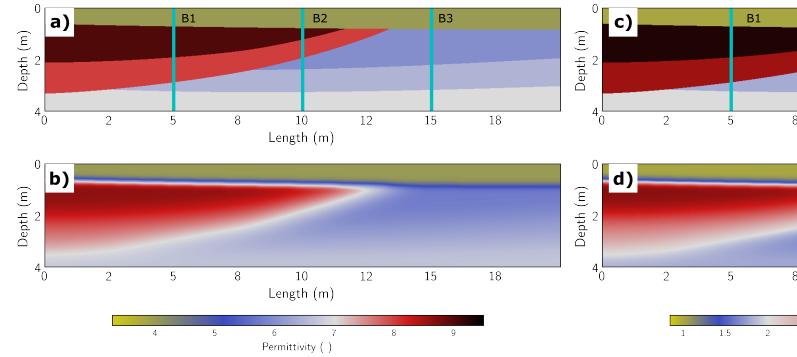


Figure 6: Synthetic alluvial aquifer true and initial parameters. True **a** and initial **b** permittivities. True **c** and initial **d** conductivities. The cyan lines represent boreholes B1, B2 and B3 from left to right.

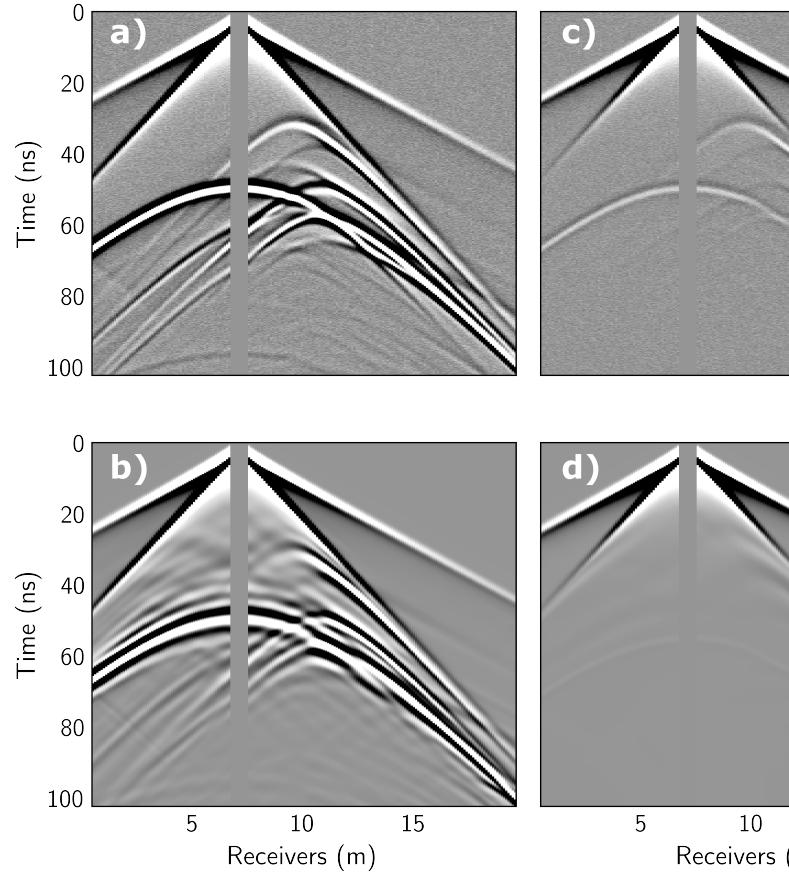


Figure 7: GPR shot gather # 7 of the low and high-conductivity scenarios and their respective best recovered parameters as given by Figures 10-11 **d** for the low-conductivity and 12-13 **d** for the high-conductivity scenario. Amplitudes are clipped to 1.5% of the maximum amplitude in the data.

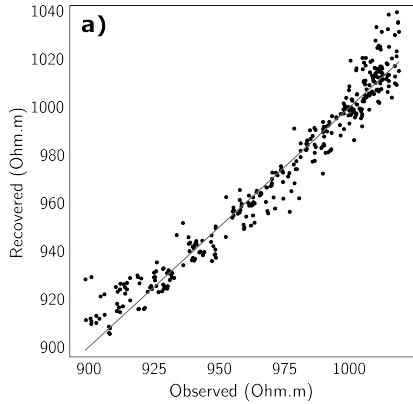


Figure 8: ER data of the low **a** and high **b** conductivity scenarios and their respective best recovered parameters.

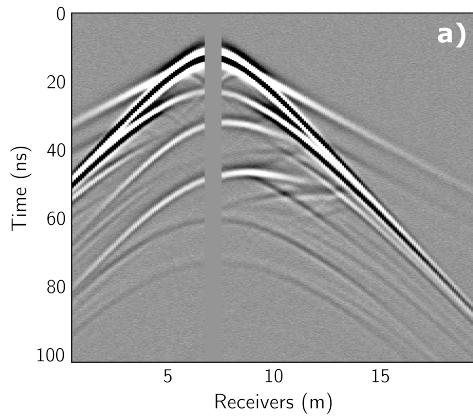
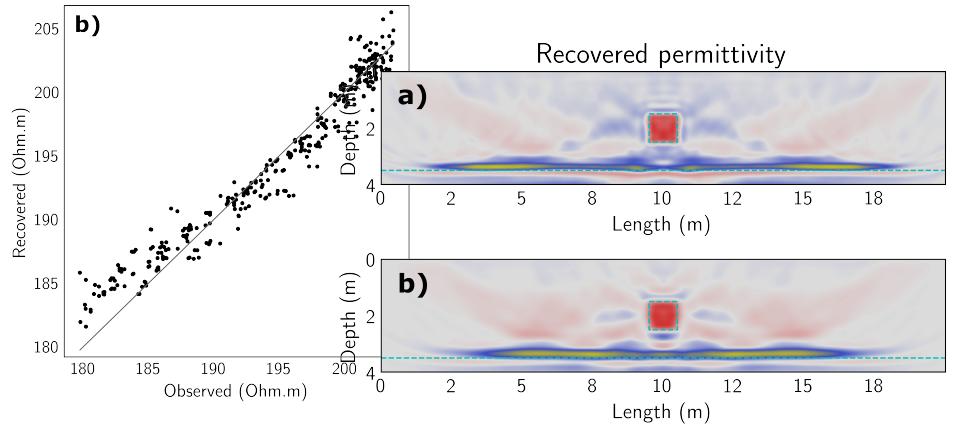
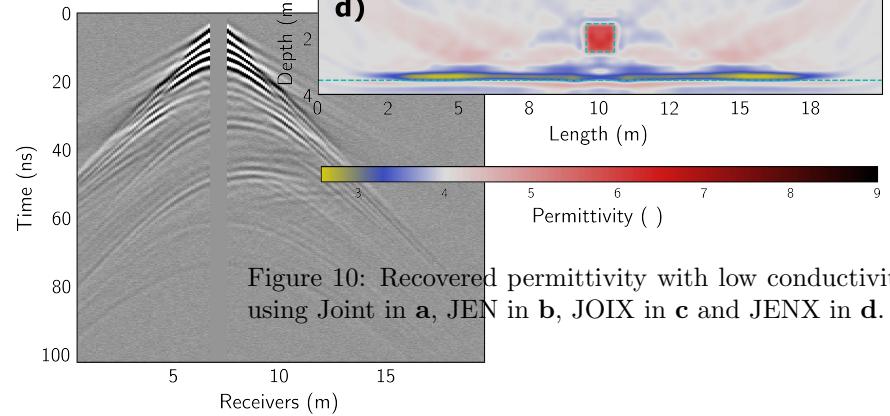


Figure 9: Residuals of GPR shot-gather #7 for the synthetic alluvial aquifer. Residual of initial model and observed in **a**, and of recovered and observed in **b**. Recovered data correspond to the JOIX method. Amplitudes are clipped to 1.5% of the maximum amplitude in the data.



Low σ	ε_r	σ	average
Joint	0.3691	0.4927	0.4309
JEN	0.3742	0.4972	0.4357
JOIX	0.3682	0.4912	0.4297
JENX	0.3697	0.4908	0.4303

	$a_{dc} \bullet$	\dot{a}_{dc}	$\dot{\Theta}_{dc}$	\dot{a}_w	$\dot{\Theta}_w$	β_{ε_r}	β_σ	h_{ε_r}	d_{ε_r}	h_σ	d_σ	iterations
Joint	0.2	3	2	1.5	0.3							
JEN	0.2	3	2	1.5	0.3	0.5	0.5					
JOIX	0.2	3	2	1.5	0.3			-10^{-3}				
JENX	0.2	3	2	1.5	0.3	0.5	0.5	-10^{-3}				

Table 4: Inversion parameters for the synthetic alluvial aquifer.

Table 5: RMS error and average of the RMS errors for all inversion methods of the true and recovered parameters for the low-conductivity scenario. The region where the errors were calculated is the band between 8 and 12 m in length. The boxed results are the smallest value of each column.

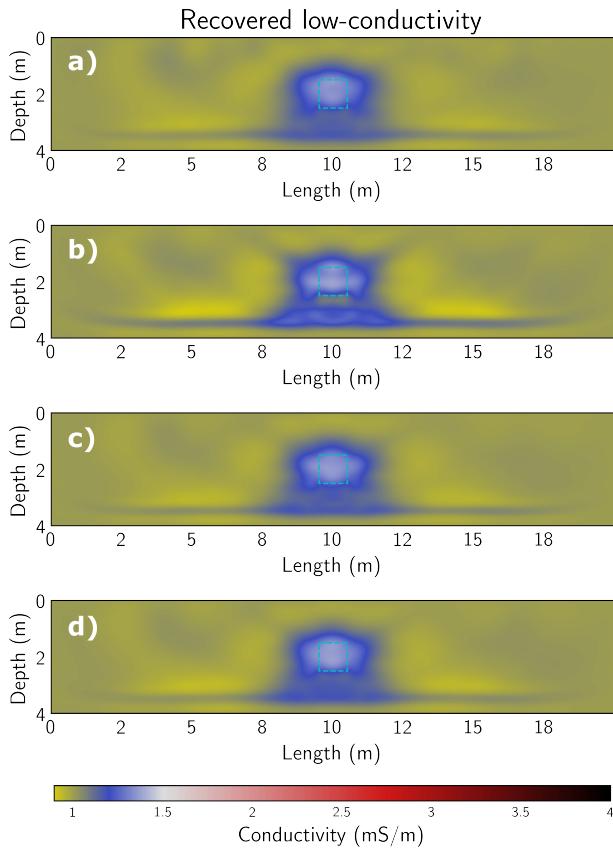


Figure 11: Recovered low conductivity using Joint in **a**, JEN in **b**, JOIX in **c** and JENX in **d**.

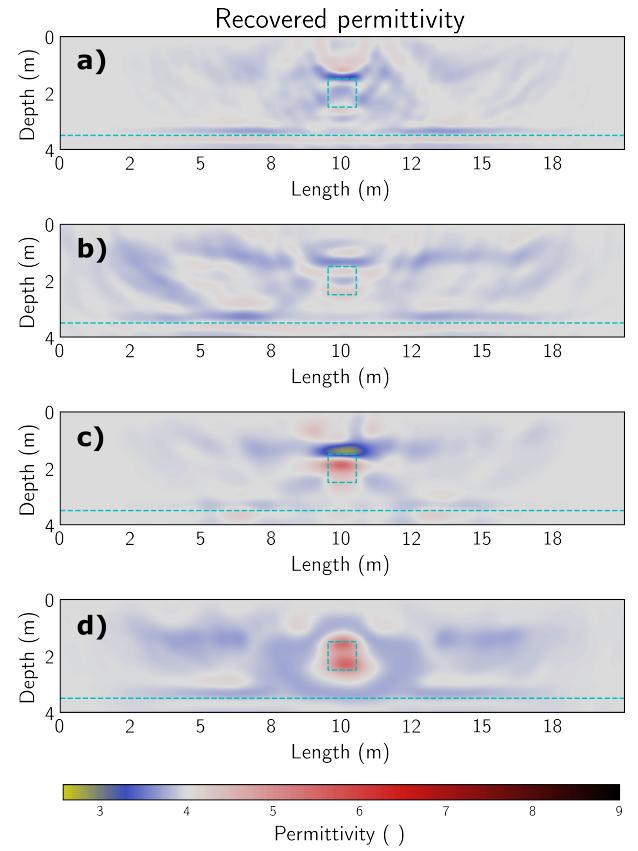


Figure 12: Recovered permittivity with high conductivity using Joint in **a**, JEN in **b**, JOIX in **c** and JENX in **d**.

High σ	ϵ_r	σ	average
Joint	0.3708	0.5012	0.4360
JEN	0.3644	0.4992	0.4318
JOIX	0.3666	0.4976	0.4321
JENX	0.3642	0.4915	0.4278

Table 6: RMS error and average of the RMS errors for all inversion methods of the true and recovered parameters for the high-conductivity scenario. The region where the errors were calculated is the band between 8 and 12 m in length. The boxed results are the smallest value of each column.

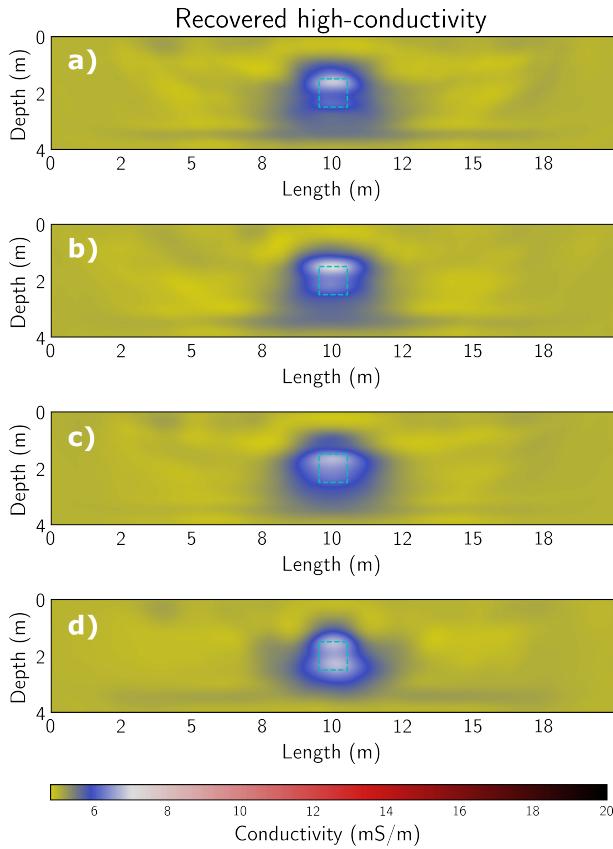


Figure 13: Recovered high conductivity using Joint in **a**, JEN in **b**, JOIX in **c** and JENX in **d**.

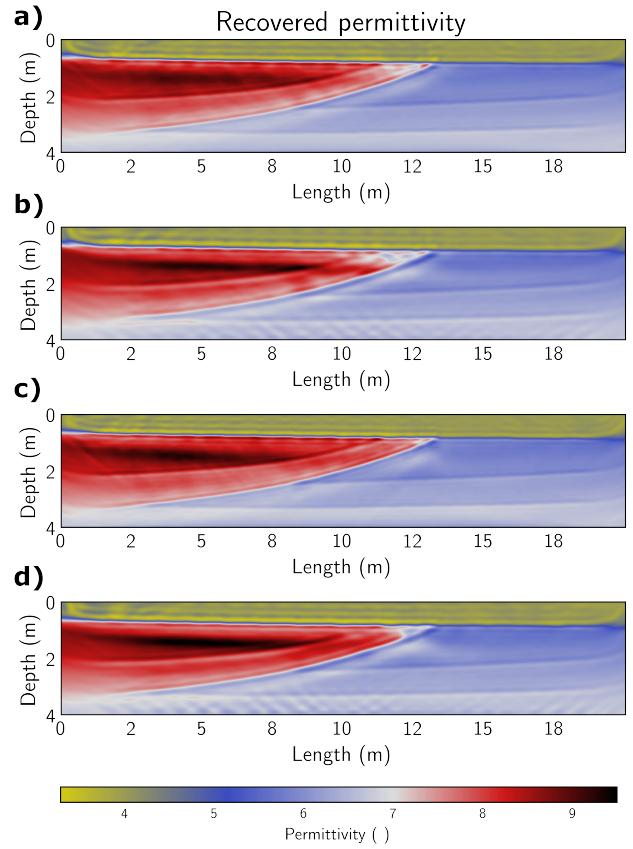


Figure 14: Recovered permittivity for the synthetic alluvial aquifer using Joint in **a**, JEN in **b**, JOIX in **c** and JENX in **d**.

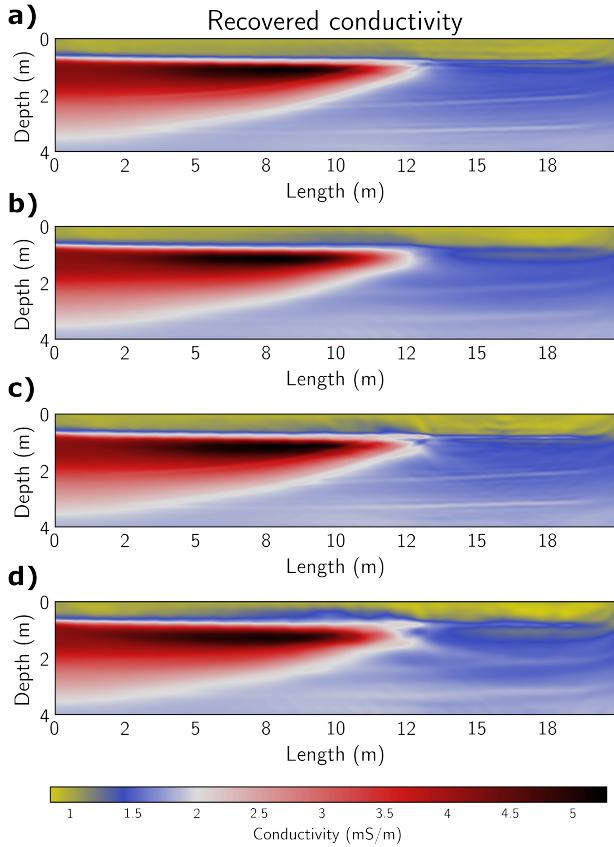


Figure 15: Recovered conductivity for the synthetic alluvial aquifer using Joint in **a**, JEN in **b**, JOIX in **c** and JENX in **d**.

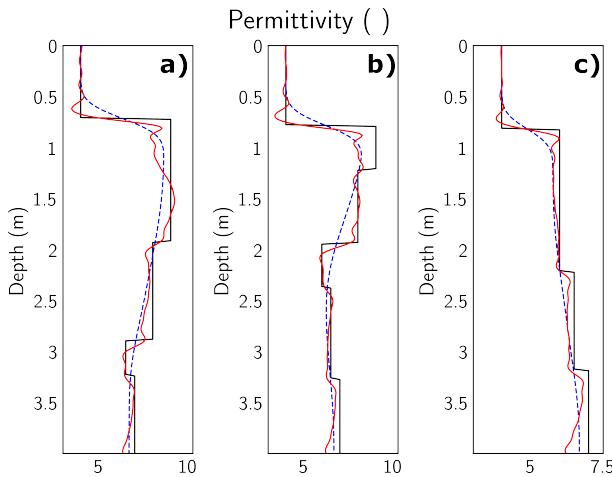


Figure 16: Recovered permittivity of the synthetic alluvial aquifer using the JOIX method on boreholes B1, B2 and B3 in **a**, **b** and **c** respectively. True is solid black and initial model is dashed blue.

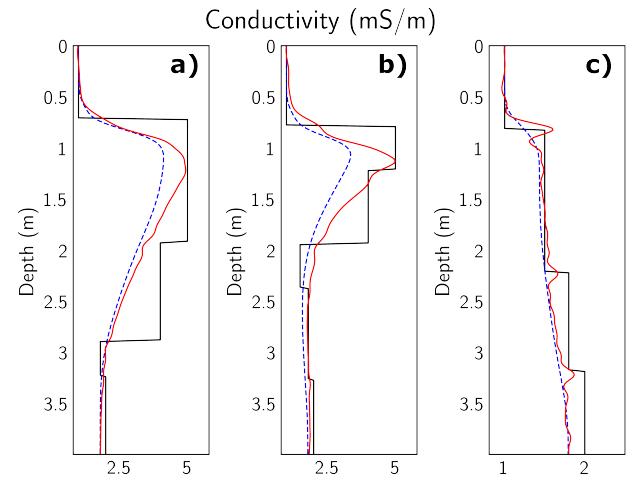


Figure 17: Recovered conductivity of the synthetic alluvial aquifer using the JOIX method on boreholes B1, B2 and B3 in **a**, **b** and **c** respectively. True is solid black and initial model is dashed blue.

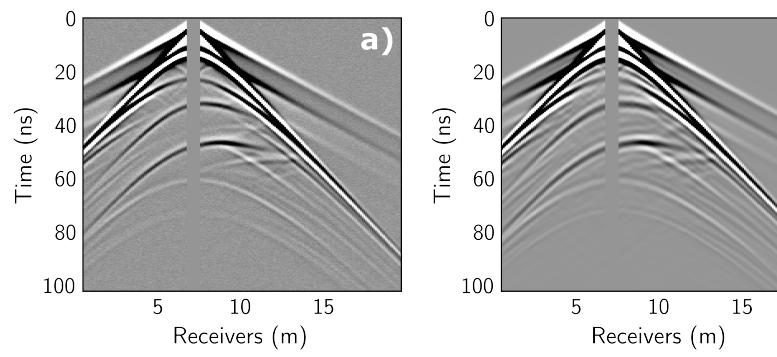


Figure 18: Synthetic alluvial aquifer data. Observed **a** and recovered **b** GPR data for shot-gather #7. In **c** observed and recovered ER data. Recovered data correspond to the JOIX method. Amplitudes are clipped to 1.5% of the maximum amplitude in the data.

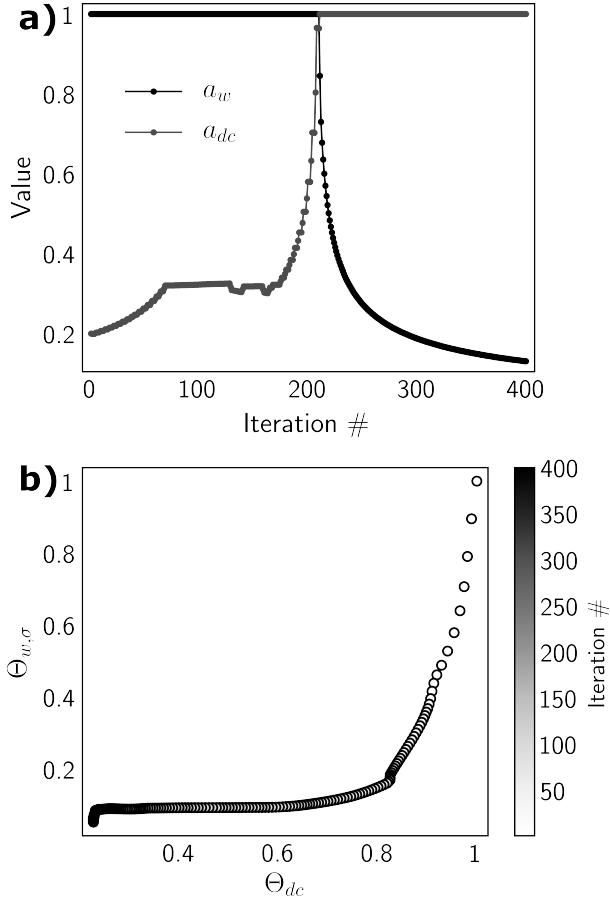


Figure 19: Inversion weights of the synthetic alluvial aquifer using the JOIX method. In **a** values of weights a_w and a_{dc} over iterations. In **b** objective function values for $\Theta_{w,\sigma}$ and Θ_{dc} .

et al. (2019) using a full waveform inversion approach,

$$\mathbf{v}_w = \mathbf{L}_w \mathbf{e}_w(-t), \quad (\text{A-1})$$

$$\mathbf{g}_{w,\sigma}^s = -\sum_t \mathbf{u}(-t) \odot \mathbf{v}_w(t) \cdot \Delta t, \quad (\text{A-2})$$

$$\mathbf{g}_{\varepsilon_r}^s = -\sum_t \dot{\mathbf{u}}(-t) \odot \mathbf{v}_w(t) \cdot \Delta t, \quad (\text{A-3})$$

where t denotes time, $(-t)$ denotes time reversed, \odot denotes element-wise multiplication, $\dot{\mathbf{u}}$ denotes the time derivative of \mathbf{u} (computed with a numerical finite-difference scheme), \mathbf{v}_w is the adjoint wavefield (the back-propagation of errors), and Δt denotes the discretized time interval.

We compute \mathbf{g}_{dc}^s using the adjoint potential field \mathbf{v}_{dc} Domenzain et al. (2019),

$$\begin{aligned} \mathbf{L}_{dc}^\top \mathbf{v}_{dc} &= \mathbf{M}_{dc}^\top \mathbf{e}_{dc}, \\ \mathbf{g}_{dc}^s &= \mathbf{S}_{dc} \mathbf{v}_{dc}, \end{aligned} \quad (\text{A-4})$$

where \mathbf{g}_{dc}^s and \mathbf{v}_{dc} are vectors of size $n_x n_z \times 1$ and $\mathbf{S}_{dc} = -((\nabla_\sigma \mathbf{L}_{dc}) \varphi)^\top$.

APPENDIX B ENVELOPE GPR GRADIENT

In order to apply the FWI scheme with the modified envelope data, we first need to deduce a new adjoint source as a result of the chain rule on our objective function. We follow Bozdag et al. (2011) and define the adjoint source of equation A-1 in the continuous case and then bring it back to the discrete case. Let u denote the y component of the electromagnetic wavefield defined in space and time for a given source. We denote the analytical representation of u by,

$$\tilde{u} = u + i\hat{u}, \quad (\text{B-1})$$

where the hat denotes the Hilbert transform of u . We will also refer to the Hilbert transform of u by $\{u\}_H$. We will modify the objective function Θ_w , and that will modify the adjoint source because of the chain rule on Θ_w .

The instantaneous amplitude of the wavefield (i.e. envelope) is,

$$u_a = \sqrt{u^2 + \hat{u}^2}. \quad (\text{B-2})$$

In what follows we will define new objective functions and find the new adjoint source for them. We will denote d_u the derivative with respect to u and use this identity derived from the definition of the Hilbert transform,

$$\int f \cdot d_u \hat{g} dt = - \int \hat{f} \cdot d_u g dt. \quad (\text{B-3})$$

Let the instantaneous amplitude objective function be,

$$\Theta_{w,a} = \frac{1}{2} \int_0^T e_{w,a}^2 dt, \quad e_{w,a} = u_a - u_a^o, \quad (\text{B-4})$$

where the superscript o denotes observed data. We need the derivative of $\Theta_{w,a}$ with respect to the parameters, and

for that we also need $d_u \Theta_{w,a}$ since u depends on the parameters. We have,

$$\begin{aligned} d_u \Theta_{w,a} &= \int_0^T e_{w,a} \cdot d_u e_{w,a} dt, \\ d_u e_{w,a} &= d_u u_a, \\ &= \frac{u + \hat{u} \cdot d_u \hat{u}}{u_a^2}. \end{aligned} \quad (\text{B-5})$$

We now invoke identity B-3 in $d_u \Theta_{w,a}$,

$$d_u \Theta_{w,a} = \int_0^T \underbrace{\frac{e_{w,a} \cdot u}{u_a} - \left\{ \frac{e_{w,a} \cdot \hat{u}}{u_a} \right\}_H}_{\text{adjoint source}} dt. \quad (\text{B-6})$$

From equation B-6 we have that in the discrete case for an observed shot-gather $\mathbf{d}_w^{o,s}$ the adjoint source for the envelope transformed data is,

$$\mathbf{s}_{w,a} = \frac{\mathbf{e}_{w,a} \cdot \mathbf{d}_w^{o,s}}{\mathbf{d}_{w,a}^{o,s}} - \left\{ \frac{\mathbf{e}_{w,a} \cdot \{\mathbf{d}_w^{o,s}\}_H}{\mathbf{d}_{w,a}^{o,s}} \right\}_H, \quad (\text{B-7})$$

where $\mathbf{d}_{w,a}^{o,s}$ denotes the envelope of the observed data and $\mathbf{e}_{w,a}$ denotes the residual of the observed envelope data and the synthetic envelope data. The gradients $\mathbf{g}_{\sigma,a}^s$ and $\mathbf{g}_{\varepsilon_r,a}^s$ are,

$$\mathbf{v}_w = \mathbf{L}_w \mathbf{s}_{w,a}(-t), \quad (\text{B-8})$$

$$\mathbf{g}_{\sigma,a}^s = - \sum_t \mathbf{u}(-t) \odot \mathbf{v}_w(t) \cdot \Delta t, \quad (\text{B-9})$$

$$\mathbf{g}_{\varepsilon_r,a}^s = - \sum_t \dot{\mathbf{u}}(-t) \odot \mathbf{v}_w(t) \cdot \Delta t. \quad (\text{B-10})$$

APPENDIX C MINIMIZING Θ_τ

We present a Gauss-Newton algorithm for optimizing Θ_τ that enables our joint inversion scheme to independently weigh the structure of $\boldsymbol{\sigma}$ over $\boldsymbol{\varepsilon}_r$ (or vice-versa).

Let \mathbf{D}_x and \mathbf{D}_z be the discretized differential operators in the x and z directions written as matrices of size $n_x n_z \times n_x n_z$,

$$\boldsymbol{\tau} = \mathbf{D}_x \boldsymbol{\varepsilon}_r \odot \mathbf{D}_z \boldsymbol{\sigma} - \mathbf{D}_z \boldsymbol{\varepsilon}_r \odot \mathbf{D}_x \boldsymbol{\sigma}. \quad (\text{C-1})$$

The derivatives of $\boldsymbol{\tau}$ with respect to $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ are,

$$\begin{aligned} \nabla_{\varepsilon} \boldsymbol{\tau} &= \mathbf{D}_x \odot [\mathbf{D}_z \boldsymbol{\sigma}] - \mathbf{D}_z \odot [\mathbf{D}_x \boldsymbol{\sigma}], \\ \nabla_{\sigma} \boldsymbol{\tau} &= \mathbf{D}_z \odot [\mathbf{D}_x \boldsymbol{\varepsilon}_r] - \mathbf{D}_x \odot [\mathbf{D}_z \boldsymbol{\varepsilon}_r], \end{aligned} \quad (\text{C-2})$$

where brackets indicate a matrix of size $n_x n_z \times n_x n_z$ and all columns of a matrix $[\mathbf{a}]$ are the column vector \mathbf{a} . Let $\mathbf{J}_{\tau,o}^\top = \nabla_o \boldsymbol{\tau}$, then the gradients of Θ_τ are,

$$\begin{aligned} \mathbf{g}_{\tau,\varepsilon} &= \mathbf{J}_{\tau,\varepsilon} \boldsymbol{\tau}, \\ \mathbf{g}_{\tau,\sigma} &= \mathbf{J}_{\tau,\sigma} \boldsymbol{\tau}. \end{aligned} \quad (\text{C-3})$$

We compute the updates of $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ by,

$$\begin{aligned} \Delta \boldsymbol{\varepsilon}_{r,\tau} &= -(\mathbf{J}_{\tau,\varepsilon} \mathbf{J}_{\tau,\varepsilon}^\top + \alpha_{\tau,\varepsilon} \mathbf{I})^{-1} \mathbf{g}_{\tau,\varepsilon}, \\ \Delta \boldsymbol{\sigma}_\tau &= -(\mathbf{J}_{\tau,\sigma} \mathbf{J}_{\tau,\sigma}^\top + \alpha_{\tau,\sigma} \mathbf{I})^{-1} \mathbf{g}_{\tau,\sigma}, \end{aligned} \quad (\text{C-4})$$

where \mathbf{I} is the identity matrix of size $n_x n_z \times n_x n_z$, and $\alpha_{\tau,\varepsilon}$ and $\alpha_{\tau,\sigma}$ are step-sizes for the optimal descent direction for the previous iteration gradients and are computed with an n-point parabola approximation. We then normalize the updates by their largest amplitude and scale them with their respective current step-sizes. At each iteration, either $\boldsymbol{\varepsilon}_r$ and $\boldsymbol{\sigma}$ are updated by,

$$\boldsymbol{\varepsilon}_r \leftarrow \boldsymbol{\varepsilon}_r + \Delta \boldsymbol{\varepsilon}_{r,\tau}, \quad (\text{C-5})$$

$$\boldsymbol{\sigma} \leftarrow \boldsymbol{\sigma} + \Delta \boldsymbol{\sigma}_\tau.$$

In order to control the weigh of either structures $\boldsymbol{\varepsilon}_r$ or $\boldsymbol{\sigma}$ in our joint inversion, at each iteration we store the update information of $\Delta \boldsymbol{\varepsilon}_{r,\tau}$ and $\Delta \boldsymbol{\sigma}_\tau$ in the *master* updates $\Delta \boldsymbol{\varepsilon}_{r,\tau,o}$ and $\Delta \boldsymbol{\sigma}_{\tau,o}$,

$$\Delta \boldsymbol{\varepsilon}_{r,\tau,o} \leftarrow \Delta \boldsymbol{\varepsilon}_{r,\tau,o} + \Delta \boldsymbol{\varepsilon}_{r,\tau}, \quad (\text{C-6})$$

$$\Delta \boldsymbol{\sigma}_{\tau,o} \leftarrow \Delta \boldsymbol{\sigma}_{\tau,o} + \Delta \boldsymbol{\sigma}_\tau. \quad (\text{C-7})$$

We note that in our inversion scheme presented in the section **Joint inversion with cross-gradients** we first optimize Θ_τ modifying $\boldsymbol{\sigma}$ and keeping $\boldsymbol{\varepsilon}_r$ fixed, and then we optimize Θ_τ modifying $\boldsymbol{\varepsilon}_r$ and keeping $\boldsymbol{\sigma}$ fixed.

APPENDIX D

INITIAL MODELS FOR THE SYNTHETIC ALLUVIAL AQUIFER

For the first initial model (see Figure D-1a), we smooth the true permittivity with a low-pass gaussian filter as to only allow two characteristic wavelengths in the space-frequency domain (a gaussian with a half-width of 0.8 1/m). For the second initial model (see Figure D-1b), we first remove the top layer from the true permittivity model, we then smooth analogously as for the first initial model, and then we return the first layer without smoothing. In order to keep the location of the shallow reflector equal in both initial permittivity and conductivity, we interpolate permittivities to obtain Figures D-1c and D-1d.

Two main differences between the first and second initial models are that the first initial model does not have an accurate amplitude in the first layer and does not follow the low velocity region in length. As a result, when compared to the inversions of the first initial model (Figures D-1e and D-1g), the second model is visibly able to resolve all layers in the model with minimal artifacts in the first layer (Figures D-1f and D-1h). We note however, that the first initial model is able to correctly identify the location of the first-second layer boundary.

We choose the initial model for the inversions presented in the main text as a perturbed true model between the two initial models presented in this Appendix. First we remove the top layer from the true permittivity model, and then smooth with a low-pass gaussian filter as to only allow two characteristic wavelengths in the space-frequency domain (a gaussian with a half-width of 0.8 1/m). Then we decrease the values by 4% of the true values, return the first layer and smooth again as to only allow six characteristic wavelengths in the space-frequency domain (a

gaussian with a half-width of 2.5 1/m). The initial model for the conductivity is achieved by interpolation of the permittivity. The result is a smooth initial model with values 4% less than the true model but with a not-so-smooth first layer interface.

Such a smooth initial velocity model can be achieved by following the inversion procedure of Bradford et al. (2009a). This method for estimating an initial velocity model is robust when air-wave refractions are present in the data, and resolves the subsurface in a top-down approach. We conclude that if the GPR field data exhibits air-wave refractions, the better the initial model fits these events in the data, the better the inversion results will be.

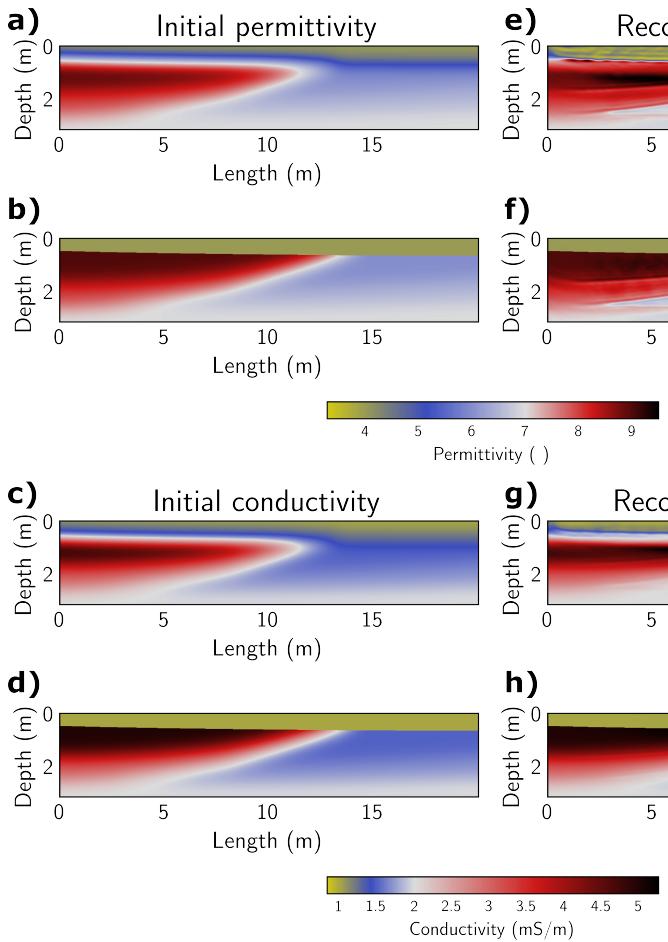


Figure D-1: Sensitivity analysis of the initial model for the synthetic alluvial aquifer. In **a**, **b**, **c** and **d** we have the first and second initial model for permittivity and conductivity. In **e**, **f**, **g** and **h** we have their respective recovered parameters by using the JOIX method.

REFERENCES

- Baeten, G., J. W. de Maag, R.-E. Plessix, R. Klaassen, T. Qureshi, M. Kleemeyer, F. t. Kroode, and Z. Ru-jie, 2013, The use of low frequencies in a full-waveform inversion and impedance inversion land seismic case study: *Geophysical Prospecting*, **61**, 701–711.
- Barrash, W., and T. Clemo, 2002, Hierarchical geostatistics and multifacies systems: Boise hydrogeophysical research site, boise, idaho: *Water Resources Research*, **38**, 14–1.
- Bozdağ, E., J. Trampert, and J. Tromp, 2011, Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements: *Geophysical Journal International*, **185**, 845–870.
- Bradford, J. H., W. P. Clement, and W. Barrash, 2009a, Estimating porosity with ground-penetrating radar reflection tomography: A controlled 3-d experiment at the boise hydrogeophysical research site: *Water Resources Research*, **45**.
- Bradford, J. H., J. T. Harper, and J. Brown, 2009b, Complex dielectric permittivity measurements from ground-penetrating radar data to estimate snow liquid water content in the pendular regime: *Water resources research*, **45**.
- Doetsch, J., N. Linde, and A. Binley, 2010, Structural inversion of time-lapse crosshole gpr and traveltime data: *Geophysical Research Letters*, **37**.
- Domenzain, D., J. Bradford, and J. Mead, 2019, Joint inversion of full-waveform gpr and er data. part 1: Inversion of time-domain crosshole gpr data: *Geophysics*, **72**, J53–J64.
- Ernst, J. R., A. G. Green, H. Maurer, and K. Holliger, 2007a, Application of a new 2d time-domain full-waveform inversion scheme to crosshole radar data: *Geophysics*, **72**, J53–J64.
- Ernst, J. R., H. Maurer, A. G. Green, and K. Holliger, 2007b, Full-waveform inversion of crosshole radar data based on 2-d finite-difference time-domain solutions of maxwell's equations: *IEEE transactions on geoscience and remote sensing*, **45**, 2807–2828.
- Flosgoso, E., and L. A. Gallardo, 2009, Cross-gradients 3d inversion with applications to gravity and magnetic data: *Geophysics*, **74**, L31–L42.
- Gallardo, L. A., and M. A. Meju, 2003, Characterization of heterogeneous near-surface materials by joint 2d inversion of dc resistivity and seismic data: *Geophysical Research Letters*, **30**.
- , 2007, Joint two-dimensional cross-gradient imaging of magnetotelluric and seismic traveltimes for structural and lithological classification: *Geophysical Journal International*, **169**, 1261–1272.
- Gross, L., 2019, Weighted cross-gradient function for joint inversion with the application to regional 3-d gravity and magnetic anomalies: *Geophysical Journal International*, **217**, 2035–2046.
- Haber, E., and D. Oldenburg, 1997, Joint inversion: a structural approach: *Inverse problems*, **13**, 63.
- Hu, W., A. Abubakar, and T. M. Habashy, 2009, Joint

- electromagnetic and seismic inversion using structural constraints: *Geophysics*, **74**, R99–R109.
- Lavoué, F., R. Brossier, L. Métivier, S. Garambois, and J. Virieux, 2014, Two-dimensional permittivity and conductivity imaging by full waveform inversion of multi-offset gpr data: A frequency-domain quasi-newton approach: *Geophysical Journal International*, **197**, 248–268.
- Linde, N., A. Binley, A. Tryggvason, L. B. Pedersen, and A. Revil, 2006, Improved hydrogeophysical characterization using joint inversion of cross-hole electrical resistance and ground-penetrating radar traveltimes: *Water Resources Research*, **42**.
- Liu, Z., and J. Zhang, 2017, Joint traveltimes and waveform envelope inversion for near-surface imaging: *Pure and Applied Geophysics*, **174**, 1269–1289.
- Meles, G. A., J. Van der Kruk, S. A. Greenhalgh, J. R. Ernst, H. Maurer, and A. G. Green, 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.
- Pidlisecky, A., E. Haber, and R. Knight, 2007, Resinv3d: A 3d resistivity inversion package: *Geophysics*, **72**, H1–H10.
- Pratt, R. G., C. Shin, and G. Hick, 1998, Gauss–newton and full newton methods in frequency–space seismic waveform inversion: *Geophysical Journal International*, **133**, 341–362.
- Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: *Geophysics*, **74**, WCC1–WCC26.