Joint Inversion of Crosshole GPR and Seismic Traveltime Data

Niklas Linde[1] and Joseph Andreas Doetsch[2]

[1]Institute of Geophysics, University of Lausanne, Lausanne, Switzerland (niklas.linde@unil.ch);

[2]Institute of Geophysics, ETH Zurich, Zurich, Switzerland (doetsch@aug.ig.erdw.ethz.ch).

ABSTRACT

Joint inversion of crosshole ground-penetrating radar and seismic data can improve model resolution and fidelity of the resultant individual models. Model coupling obtained by minimizing or penalizing some measure of structural dissimilarity between models appears to be the most versatile approach, since only weak assumptions about petrophysical relationships are required. Nevertheless, experimental results and petrophysical arguments suggest that when porosity variations are weak in saturated unconsolidated environments then radar wavespeed is approximately linearly related to seismic wavespeed. Under such circumstances, model coupling can also be achieved by incorporating cross-covariances in the model regularization. We present two case studies in which structural similarity is imposed by penalizing models for which the model cross-gradients are non-zero. The first case study demonstrates improvements in model resolution by comparing the resulting models with borehole information, whereas the second uses point-spread functions. Although, radar-seismic wavespeed crossplots are found to be very similar for the two case studies, the models plot in different portions of the graph, suggesting differences in porosity. Both examples display a close quasi-linear relationship between radar-seismic wavespeed in unconsolidated environments that is rather well described by the corresponding lower Hashin-Shtrikman

bounds. We suggest that combining crossplots of the joint inversion models with Hashin-Shtrikman bounds can better constrain porosity and pore structure than individual inversion results.

INTRODUCTION

Joint inversions of geophysical data can:

1. improve model resolution and fidelity of individual models;

2. provide consistent geophysical models for interpretation, classification and petrophysical inference;

3. make it easier to identify modeling and geometrical errors by comparing the models obtained by individual and joint inversions;

4. allow hypotheses testing concerning geological structure, processes, and petrophysical relationships.

Numerous methodologies to jointly invert disparate but co-located geophysical data at different scales and for different applications have been developed and tested in the last decades (e.g., Vozoff and Jupp, 1975; Lines et al., 1988; Tryggvason et al., 2002; Gallardo and Meju, 2003; Musil et al., 2003; Monteiro Santos et al., 2006). Many critical choices for the development of joint and individual inversion algorithms are similar. These choices relate to model parameterization, model regularization, model and data norm, type of forward models and equation solvers, and stochastic versus deterministic frameworks. Difficulties related to weighting different data sets (e.g., Lines et al., 1988) are not so different from difficulties that arise when inverting single geophysical data types (e.g., should one assume absolute or relative errors or a mixture of the two? how are actual errors estimated? etc.). Data weighting for joint inversion needs to not only consider data and modeling errors, but also sensitivity with respect to the model parameters of interest and data redundancy that arise

when many data points provide very similar information. The fundamental difference between joint and individual inversion is the need to couple the models at the inversion stage. There are basically four different approaches for doing this:

1. a structural approach in which it is assumed that models share one or several boundaries or that some measure of model structure is similar over given model domains (e.g., Haber and Oldenburg, 1997; Gallardo and Meju, 2003);

2. an approach whereby models are explicitly linked with known or unknown (i.e., to be determined during the inversion) petrophysical relationships in order to create as many inversion models as there are data sets; one example would be the joint inversion of P- and S-wave traveltimes in which joint inversion is achieved by damping the models against a predefined V_p/V_s ratio (e.g., Tryggvason et al., 2002);

3. an approach in which the joint inverse problem is formulated in terms of one inversion parameter only that is considered of primary importance, whereas the other data sources provide "proxy" data related to this primary inversion parameter through petrophysical relationships; typical examples in the field of hydrogeophysics would be using measurements of the hydrological state in boreholes (salinity, pressure, or water content) in response to hydrological testing together with crosshole geophysical data (also sensitive to these state variables) to invert directly for the permeability structure (e.g., Kowalsky et al., 2005);

4. an approach in which model parameters correspond to properties that are only indirectly related to the geophysical data at hand (i.e., no partial differential equations that describe the physical system are available); an example would be to invert for the spatially distributed electrical formation factor and surface conductivity using radar

traveltimes and attenuation data; other examples could include inversion for the geochemical composition using diverse geophysical data (e.g., Chen et al., 2004).

Structural approaches (Category 1 above) provide robust solutions for a wide range of application types in deterministic joint inversion. We focus here on joint inversion based on the cross-gradients constraints introduced by Gallardo and Meju (2003, 2004). A number of other interesting structural approaches have been presented in the literature. For example, Hyndman and Harris (1996) present a traveltime inversion scheme for inverting two-dimensional zonal models using crosshole seismic traveltime data. Their technique could easily be extended to joint inversion of radar and seismic traveltime and attenuation data by assuming that all these data sets are sensitive to the same uniform zones and zonal boundaries. Paasche and Tronicke (2007) and Paasche et al. (2008) present an iterative sequential approach to invert crosshole radar traveltime and attenuation data. Their approach combines gradient-based deterministic inversion with a cluster algorithm that is used after each iteration step to classify the models in terms of a number of zones. This zonal model is used as the starting model for the next iteration step.

The examples presented here focus on the joint inversion of crosshole radar and seismic traveltimes. It would be rather straight-forward to modify the algorithm presented here to accommodate radar attenuation data (Holliger et al., 2001), Fresnel volume inversion (Vasco et al., 1995), full-waveform inversion (e.g., Ernst et al., 2007; Belina et al., 2009; Pratt, 1999), or for joint inversion of surface-based seismic refraction (e.g., Lanz et al., 1998) and GPR reflection data (e.g., Bradford et al., 2009).

In this chapter, the joint inversion methodology is first introduced before two case-studies are presented together with a discussion about cross-property relations of seismic and radar wavespeeds. The chapter ends with discussion and conclusions.

METHOD

Joint inversion based on structural coupling using the cross-gradients constraints was introduced by Gallardo and Meju (2003; 2004). This approach has been adapted and applied to a wide range of data types (Gallardo and Meju, 2003, 2004, 2007; Gallardo et al., 2005; Gallardo, 2007; Linde et al., 2006, 2008; Tryggvason and Linde, 2006; Fregoso and Gallardo, 2009; Hu et al., 2009; Doetsch et al., submitted). The normalized cross-gradients function $\mathbf{t}_{qr}{}'(x,y,z)$ of two models $\mathbf{m}_q$ and $\mathbf{m}_r$ at location x, y, z is (Linde et al., 2008):

$$\mathbf{t}_{qr}{}'(x,y,z) = \frac{\nabla\mathbf{m}_q(x,y,z) \times \nabla\mathbf{m}_r(x,y,z)}{\left|m_q(x,y,z)\right| \cdot \left|m_r(x,y,z)\right|}, \tag{1}$$

where $\nabla\mathbf{m}_q(x,y,z)$ and $\nabla\mathbf{m}_r(x,y,z)$ are the gradients of models $\mathbf{m}_q$ and $\mathbf{m}_r$ at location x, y, z. The original definition of the cross-gradients function $\mathbf{t}_{qr}(x,y,z)$ by Gallardo and Meju (2003) does not include the normalization term, which facilitates comparison of results from different applications and different joint inversion implementations. Constraints based on the cross-gradients function allows one of the models to change at a given position without requiring the other to change and it focuses on the direction of the change rather than the magnitude. The cross-gradients function is typically discretized using forward (e.g., Gallardo and Meju, 2003) or central differences (e.g., Linde et al., 2008). The discretized cross-gradients function based on central-differences for the y-component $t_{qr}^y(i,j,k)$ for two models $\mathbf{m}_q$ and $\mathbf{m}_r$ with a uniform discretization is:

$$t_{qr}^{y}(i,j,k) = \frac{1}{4\Delta x \Delta z}\left[m_q(i,j,k+1) - m_q(i,j,k-1)\right]\left[m_r(i+1,j,k) - m_r(i-1,j,k)\right]$$
$$-\frac{1}{4\Delta x \Delta z}\left[m_q(i+1,j,k) - m_q(i-1,j,k)\right]\left[m_r(i,j,k+1) - m_r(i,j,k-1)\right],$$

(2)

where Δx, and Δz are the discretizations in the x- and z-directions, and indices i, j, and k indicate the corresponding indices of the model cells. Gallardo and Meju (2004) provide a formulation for non-uniform cell-spacings with rectangular cells.

The cross-gradients function can either be defined for the total model (Gallardo and Meju, 2003) or for the model update with respect to a reference model (Tryggvason and Linde, 2006). The latter definition is useful when including seismic data, since there might be strong vertical trends in seismic wavespeed that dominate any effects due to small-scale variations in lithology.

The cross-gradients function is non-linear, such that it is necessary to linearize it when performing deterministic inversions. This means that an iterative approach is needed, even when solving linear forward problems (e.g., when ray-paths are assumed to be straight). Cross-gradients constraints add further non-linearity to already non-linear problems. This makes it even more important than for individual inversions to ensure a slow convergence to create final models with the least artifacts possible. The visual aspects of the joint inversion models are not very different when obtained using five or twenty iterations to achieve the target data misfit, but smaller details appear in the scatter plots of the two models and the resulting magnitude of the cross-gradients function is smaller when using many iterations. Thus, satisfactory results can be obtained using the same number of iterations as for the individual inversions, but the results are slightly improved when using more iterations, which is not a constraint for computationally benign crosshole traveltime tomography applications.

The non-linearity of the cross-gradients function makes the choice of either treating the cross-gradients as hard constraints (Gallardo and Meju, 2003, 2004) or soft constraints (Tryggvason and Linde, 2006) a matter of convenience, with no significant influence on the resulting models. Linde et al. (2008) suggests that the most important factor to effectively minimize the cross-gradients constraints is to ensure slow convergence (i.e., small model updates) during the inversion. Hu et al. (2009) solve a joint inverse problem in an iterative sequential manner in which the cross-gradients constraints are applied with respect to one model that remains fixed and one model that is updated. This approach decreases computation time and the non-linearity at each iteration step, but no comparison has been made with results obtained by simultaneous model updates. These authors also improve the convergence by using a Gauss-Newton method (i.e., second-order Taylor expansion of the objective function compared to first-order Taylor expansions in previous work).

When performing joint inversion of geophysical data with cross-gradients constraints, the objective function Φ is:

$$\Phi = \Phi_{\mathrm{d}} + \Phi_{\mathrm{m}} + \Phi_{\mathrm{CG}}, \tag{3}$$

where Φ_{d} is a data misfit term, Φ_{m} is a model structure term, and Φ_{CG} is a structural dissimilarity term as defined by the cross-gradients function. Φ_{d} is given by

$$\Phi_{\mathrm{d}} = \sum_{q=1}^{Q} \left[\mathbf{C}_{\mathrm{d},q}^{-0.5} \left(\mathbf{d}_q - \mathbf{F}_q \left(\mathbf{m}_q \right) \right) \right]_p - \Phi_{\mathrm{d}}^{*}, \tag{4}$$

where Q is the number of data types, $\mathbf{C}_{\mathrm{d},q}$ is the data error covariance matrix for model q (typically assumed to be a diagonal matrix), $\mathbf{d}_q$ are the observed data for data type q, $\mathbf{F}_q \left(\mathbf{m}_q \right)$

is the forward response of model q, and Φ_d^* refers to the pre-defined target data misfit. The forward model usually needs a finer discretization than that used for the inversion, which makes it necessary to interpolate $\mathbf{m}_q$ on to a finer grid to solve the forward problem accurately. A key problem for any inversion strategy is to obtain an accurate representation of $\mathbf{C}_{\mathrm{d},q}$ and to make a good choice of Φ_d^*. This problem is not specific to joint inversion and it will not be discussed any further here.

The norm p in Equation 4 is typically 2, which assumes that a Gaussian distribution with zero mean is assumed to characterize the data noise. To decrease the sensitivity to outliers or fat tails in the data error distribution it is useful to work with approximations of the l_p-norms for the case when $p=1$ by using iteratively reweighted least squares (IRLS) (e.g., Farquharson, 2008). The l_p-norm is given by

$$\|\mathbf{r}\|_p^p = \sum_{n=1}^{N} |r_n|^p , \tag{5}$$

where the entries in r_n denotes data residuals. In practice, one uses a normal least-squares inversion but with a matrix that re-weights $\mathbf{C}_{\mathrm{d},q}^{-0.5}$ by multiplying it with a diagonal matrix $\mathbf{R}_{\mathrm{d},q}$ with elements

$$R_{nn} = \sqrt{\left[p\left(\left(r_n\right)^2 + \gamma^2\right)^{p/2-1} \right]} , \tag{6}$$

where r_n is the corresponding data residual at the previous iteration. To approximate an l_1-norm it is common to use $p = 1$ and $\gamma = 0.1$ (Farquharson, 2008) . This re-weighting yields convergence characteristics similar to those of quadratic functions, while being almost as robust to outliers as l_1-norm inversions.

Φ_m is given by

$$\Phi_{\mathrm{m}} = \sum_{q=1}^{Q} \varepsilon_q \left[\mathbf{C}_{\mathrm{m},q}^{-0.5} \left(\mathbf{m}_q - \mathbf{m}_q^{ref} \right) \right]_p, \tag{7}$$

where ε_q acts as a trade-off parameter between data fit and model roughness for model q (the value of ε_q takes in our implementation the same value ε for all data sources and is progressively lowered at each iteration by, for example, 10-50% until the target data misfit Φ_{d}^* is reached). $\mathbf{C}_{\mathrm{m},q}$ is the corresponding model covariance matrix, and $\mathbf{m}_q^{ref}$ is the reference model for data type q.

The variance of $\mathbf{C}_{m,q}$ is often not known precisely and can thus be used to tune the individual inversions such that they reach the target data misfit at the same value of ε_q for each data type. We will see later that this is important to avoid too many tuning parameters when performing the joint inversion. In practice, $\mathbf{C}_{\mathrm{m},q}$ is often replaced by damping and smoothness constraints (Maurer et al., 1998). Damping is unsuitable for joint inversion with cross-gradients constraints, because these regularization operators have no spatial support. Instead, isotropic (Gallardo and Meju, 2003) and anisotropic (Linde et al., 2006) smoothness constraints have been used. We have found that stochastic regularization operators (see Appendix A) as introduced by Linde et al. (2006a) generally outperforms smoothness constraints in terms of convergence, stability, and resultant models that better correspond to complementary ground truth measurements (Linde et al., 2008). These operators are typically based on an exponential covariance function with integral scales that specify the spatial correlation in each direction. The stochastic regularization operator can either be based on geostatistical analysis of geophysical logging data (Linde et al., 2006a, 2008) or estimated resolution properties of the inverse problem (Doetsch et al., submitted). Stochastic regularization operators define a physical length scale that can be related to field conditions, such that fine-tuning is not needed when changing the model discretization from an initially

coarse to a more finely discretized inversion grid. Note that if it is known that two model properties have a strong linear correlation it is also possible to make a joint inversion for these properties by including additional smoothness constraints that operate between model parameters at the same location as suggested by Gallardo and Meju (2004).

Inversion results generally improve with the quality of $\mathbf{m}_q^{ref}$. In settings in which the largest variability occurs in the vertical direction (i.e., groundwater table, sediment-bedrock interface, sedimentary layers), it is beneficial to use average zero-offset profiles to define a one-dimensional $\mathbf{m}_q^{ref}$. To resolve sharper features it can help to work with IRLS mimicking l_1-norms (defined in an analogous manner as for the data misfit in Equations 4-6) instead of the traditional l_2-norm. Other approaches based on iterative reweighting (e.g., Zhdanov, 2009) may be valuable in traveltime tomography (Ajo-Franklin et al., 2007).

The last component Φ_{CG} in the objective function assures coupling between the models and is given by

$$\Phi_{\mathrm{CG}} = \sum_{q=1}^{Q} \sum_{r>q}^{R} \lambda_{qr} \left[\mathbf{T'}_{qr} \left(\mathbf{m}_q, \mathbf{m}_r \right) \right]_p , \qquad (8)$$

where λ_{qr} is a constant weight given to the cross-gradients constraints between two model types q and r, and $\mathbf{T'}_{qr} \left(\mathbf{m}_q, \mathbf{m}_r \right)$ is a vector that consists of the estimated normalized cross-gradients function (see Equation 1) in all directions and at all locations were structural similarity is imposed. Linde et al. (2008) explored sensitivities related to the choice of λ_{qr}. They determined that it can be chosen on the basis of trial inversions in which λ_{qr} varies over several orders of magnitude with one or two values of λ_{qr} for each order of magnitude. The value chosen is the one for which the mean value of $\mathbf{T'}_{qr} \left(\mathbf{m}_q, \mathbf{m}_r \right)$ is the smallest when Φ_{d}^* is

reached. When jointly inverting three data sets, Doetsch et al. (submitted) chose to assign $\lambda=\lambda_{12}=\lambda_{13}=\lambda_{23}$. The number of constraints in Equation 8 becomes impractical in three-dimensions when jointly inverting more than three data sets. Gallardo (2007) presented an alternative formulation by introducing a reference gradient defined as the strongest model gradient at each location in space.

An iterative solution of the joint inversion problem is needed because:

1. the forward responses typically vary non-linearly with the model (e.g., rays bend in heterogeneous media);

2. the cross-gradients function is non-linear, involving the product of two model gradients;

3. IRLS and other compact regularization operators use iterative reweighting.

The estimated forward response $\mathbf{d}_q^{l+1}$ of model $\mathbf{m}_q^{l+1}$ at iteration $l+1$ is given by:

$$\mathbf{d}_q^{l+1} = \mathbf{F}_q\left(\mathbf{m}_q^l\right) + \mathbf{J}_q^l \Delta\mathbf{m}_q^{l+1}, \tag{9}$$

where $\mathbf{J}_q^l$ is the Jacobian evaluated for model $\mathbf{m}_q^l$, $\mathbf{F}_q\left(\mathbf{m}_q^l\right)$ is the forward response of this model and $\Delta\mathbf{m}_q^l$ is a proposed model update. In traveltime tomography in which the slowness structure is represented by cells of constant slowness, the elements of the Jacobian are the ray length within each cell. Below we describe how we obtain $\Delta\mathbf{m}_q^{l+1}$.

Linearization of the cross-gradients function is given here for $\mathbf{t}_y'^{l+1}$ (see Equations 1 and 2)

$$\mathbf{t}_y'^{l+1} \cong \mathbf{t}_y'^l + \mathbf{B}_y^l \begin{pmatrix} \Delta\mathbf{m}_1^{l+1} \\ \Delta\mathbf{m}_2^{l+1} \end{pmatrix}, \tag{10}$$

where $\mathbf{B}_y^l$ is the Jacobian of the normalized cross-gradients function in the y-direction (Equation 2) with respect to the model parameters. Extensions of the joint inversion framework to three or more methods are straight-forward (Gallardo, 2007; Doetsch, et al., submitted) but we focus here on two methods for simplicity. At each iteration, we solve the following system of equations in a least-squares sense:

$$
\begin{bmatrix}
\mathbf{R}_d^l\left(\mathbf{C}_d\right)^{-0.5}\mathbf{J}^l \\
\varepsilon^l\mathbf{R}_m^l\mathbf{C}_m^{-0.5} \\
\lambda\mathbf{B}_x^l \\
\lambda\mathbf{B}_y^l \\
\lambda\mathbf{B}_z^l
\end{bmatrix}
\begin{bmatrix}
\Delta\mathbf{m}^{l+1}
\end{bmatrix}
=
\begin{bmatrix}
\mathbf{R}_d^l\left(\mathbf{C}_d\right)^{-0.5}\left(\mathbf{d}-\mathbf{F}\left(\mathbf{m}^l\right)\right) \\
\varepsilon^l\mathbf{R}_m^l\mathbf{C}_m^{-0.5}\left(\mathbf{m}^{ref}-\mathbf{m}^l\right) \\
-\lambda\mathbf{t}_x^{\prime l} \\
-\lambda\mathbf{t}_y^{\prime l} \\
-\lambda\mathbf{t}_z^{\prime l}
\end{bmatrix},
\tag{11}
$$

where

$$
\Delta\mathbf{m}^{l+1}=
\begin{bmatrix}
\Delta\mathbf{m}_1^{l+1} \\
\Delta\mathbf{m}_2^{l+1}
\end{bmatrix},\quad
\mathbf{m}^l=
\begin{bmatrix}
\mathbf{m}_1^l \\
\mathbf{m}_2^l
\end{bmatrix},\quad
\mathbf{m}^{ref}=
\begin{bmatrix}
\mathbf{m}_1^{ref} \\
\mathbf{m}_2^{ref}
\end{bmatrix},
$$

$$
\mathbf{d}=
\begin{bmatrix}
\mathbf{d}_1 \\
\mathbf{d}_2
\end{bmatrix},\quad
\mathbf{F}\left(\mathbf{m}^l\right)=
\begin{bmatrix}
\mathbf{F}_1\left(\mathbf{m}_1^l\right) \\
\mathbf{F}_2\left(\mathbf{m}_2^l\right)
\end{bmatrix},\quad
\mathbf{J}^l=
\begin{bmatrix}
\mathbf{J}_1^l \\
\mathbf{J}_2^l
\end{bmatrix},
$$

$$
\mathbf{C}_d^{-0.5}=
\begin{bmatrix}
w_1\mathbf{C}_{d,1}^{-0.5} & 0 \\
0 & w_2\mathbf{C}_{d,2}^{-0.5}
\end{bmatrix},\quad
\mathbf{C}_m^{-0.5}=
\begin{bmatrix}
w_1\mathbf{C}_{m,1}^{-0.5} & 0 \\
0 & w_2\mathbf{C}_{m,2}^{-0.5}
\end{bmatrix},
$$

$$
\mathbf{R}_d^l=
\begin{bmatrix}
\mathbf{R}_{d,1}^l & 0 \\
0 & \mathbf{R}_{d,2}^l
\end{bmatrix},\quad
\mathbf{R}_m^l=
\begin{bmatrix}
\mathbf{R}_{m,1}^l & 0 \\
0 & \mathbf{R}_{m,2}^l
\end{bmatrix},
$$

where w_1 and w_2 are the weights given to each data and corresponding model type in the inversion. Recall that the weight given to the cross-gradients constraints λ is constant during the inversion and that the variances of $\mathbf{C}_{m,1}$ and $\mathbf{C}_{m,2}$ are determined from the individual inversions such that the same normalized data misfit is obtained for the same value of ε_1^l and

ε_2^l. It is very important that the final models obtained by the individual and joint inversions have comparable data fit (within a few percent) to make it possible to assess the possible benefits of joint inversions. This objective explains the need for the weights w_1 and w_2. To ensure that a similar importance is given to each model, we make the first inversions with weights w_1 and w_2 that are inversely proportional to the number of data of each data type. The weight w_2 is then manually adjusted typically in the range ±30% to ensure that the final models have a similar target data misfit (Doetsch et al., submitted).

The resulting system of equations is stored as a sparse matrix and is, at each iteration, solved with the conjugate gradient method LSQR (Paige and Saunders, 1982), which has the advantage that the original condition number of Equation 8 is preserved. A preconditioner is applied that ensures that the l_2-norm of each column in the left-hand side of Equation (11) is unity, which avoids unnecessary ill-conditioning (Paige and Saunders, 1982).

RESULTS

Oyster Case-Study

We now discuss the two-dimensional joint inversion of radar and seismic data acquired between wells S14 and M3 at the South Oyster Focus Area, Virginia (Hubbard et al., 2001; Linde et al., 2008). These data sets were originally acquired to construct a permeability field to evaluate the role of heterogeneities in controlling the field-scale transport of bacteria injected for remediation purposes. The geology comprises rather coarse and high-porosity marine shoreface deposits. Radar data were acquired using a PulseEKKO 100 system with 100-MHz nominal-frequency antennae and a transmitter and receiver spacing of 0.125 m in each borehole. Seismic data were acquired using a Geometrics Strataview seismic system, a Lawrence Berkeley National Laboratory piezoelectric source, and an ITI string of hydrophone sensors, and a 0.125 m source and sensor spacing. The source pulse had a center frequency of

4 kHz, with a bandwidth of approximately 1 - 7 kHz. From these data sets, 3248 radar and 2530 seismic traveltimes were extracted.

We used a cell-discretization of 0.125 m × 0.125 m for our forward modeling 0.25 m × 0.25 m for the inversion. All tomographic inversions were stopped once the target data misfits of 0.5 ns (radar) and 20 µs (seismic) were reached. Stochastic regularization based on an exponential model (see Appendix A; Deutsch and Journel, 1998) with vertical and horizontal integral scales of 0.28 and 1.4 m were employed (Hubbard et al., 2001). The traveltimes and Jacobians were calculated in the high-frequency limit (Podvin and Lecomte, 1991; Tryggvason and Bergman, 2006) using pstomo_eq (Tryggvason et al., 2002).

The individually inverted radar (Figure 1a) and seismic (Figure 1c) tomograms display predominantly layered structures with small velocity variations and overall low velocities, diagnostic of high porosity unconsolidated sediments. The joint inversion tomograms (Figure 1b and 1d) models display slightly more distinct boundaries between facies, but the overall structure is similar to the individually inverted results. Comparison of the cross-gradients function for the individually and jointly inverted data (Figure 1e and 1f) demonstrate that the joint inversion has decreased the cross-gradients function by more than two orders of magnitude. Differences between the individual and joint inversion results are best represented by scatter plots of the seismic and radar wavespeeds (Figure 1g and 1h). Note the much higher scatter of the individual inversion wavespeeds (Figure 1g) vis-à-vis the joint inversion values (Figure 1h).

To determine if the joint inversion models provide a better representation of subsurface architecture than the individual inversion ones, Linde et al. (2008) compared the models in the vicinity of the right borehole with hydraulic conductivity estimates based on flowmeter measurements and a pumping test (Figure 2a). Trends of the co-located radar wavespeed (Figure 2b) and seismic wavespeed (Figure 2c) are very similar to the hydraulic conductivity

pattern. Correlation coefficients between log hydraulic conductivity and radar wavespeed are 0.72 and 0.78 for the individual and joint inversions. Corresponding values are 0.60 and 0.69 for the seismic wavespeed.

This case study demonstrates that joint inversion of crosshole radar and seismic traveltime data somewhat improves resolution compared to individual inversion, thus yielding an improved hydrogeophysical characterization of the investigation site.

Thur River Case-Study

Our second case study involves three-dimensional joint inversion of radar and seismic traveltime data acquired in the vicinity of the Thur River, Northern Switzerland (Doetsch et al., submitted). These data sets were acquired to delineate the main hydrostratigraphic sub-units of a gravel aquifer. The resulting models will be used in an ongoing high-resolution hydrogeophysical study aiming at improving our understanding of groundwater-river water interactions in Alpine Valleys. The geology is composed of coarse gravelly river deposits with a rather wide grain-size distribution that includes small fractions of fines (Diem et al., submitted).

Crosshole radar data at a 0.4 ns sampling rate were acquired using a RAMAC 250 MHz system, which at the site had a center frequency of about 100 MHz with energy in the 50-170 MHz frequency range. A sparker source was used to generate seismic waves with a center frequency of about 1 kHz, and a Geometrics GEODE system and a hydrophone streamer were used to record the seismic data at a sampling rate of 21 μs. Borehole deviations were measured with a deviation probe using a 3-axis fluxgate magnetometer for bearing and a 3-axis accelerometer for inclination.

We inverted the crosshole radar and seismic data acquired between four boreholes located at the corners of a 5 m × 5 m square, approximately 10 m from the Thur River. These

data were acquired across all 6 planes between the four boreholes over the 6 m thick depth interval that constituted the saturated part of the aquifer. Seismic data were recorded using source and receiver spacings of 0.25 m, whereas the radar data were collected with source and receiver spacings of 0.5 m and 0.1 m, respectively. To ensure symmetric radar coverage, the source and receiver antennas were interchanged and the experiments repeated for each plane. A total of 2661 seismic and 5584 radar traveltimes could be reliably picked (radar traveltimes affected by refractions at the groundwater table were discarded). Examples of the raw data are given in Figure 3.

A cell-discretization of 0.0625 m × 0.0625 m was employed for the forward modeling and 0.25 m × 0.25 m for the inversion. Target data misfits corresponding to a relative error of 1% for both the radar and seismic traveltimes were estimated from reciprocal measurements. All tomographic inversions were stopped once the target misfits were reached. The stochastic regularization was based on an exponential model (see Appendix A; Deutsch and Journel, 1998) with vertical and horizontal integral scale of 0.75 m and 1.5 m. This choice of weak anisotropy was made to qualitatively honor the subsurface layering seen in the borehole cores without imposing excessive lateral constraints. The integral scales were chosen in a pragmatic manner to be comparable to the resolving capabilities of the geophysical data but smaller than the borehole spacing.

The main advantage of performing ray-based three-dimensional inversion of traveltime data at the site compared with a series of six two-dimensional inversions of the data acquired along each tomographic plane is that the regions close to the four boreholes are better resolved and that the corresponding models are internally consistent at the borehole locations. The additional constraints offered by the three-dimensional inversion in the near-borehole region also help to improve the models in-between the boreholes. Any isolated anomalies located away from the tomographic plane will neither be resolved in the two- nor in the three-

dimensional inversion. The models obtained from the three-dimensional inversion in regions in-between the planes should rather be viewed as interpolations between the models along the planes using the stochastic regularization operator.

Both individual seismic (Figure 4a) and radar (Figure 4b) inversions resolve a centrally located high-velocity zone imbedded in a background of lower velocities. A very similar model was also obtained from inversion of crosshole geoelectric data; the high-velocity zone shows up as a region of low resistivity (Doetsch et al., in press; Doetsch et al., submitted). The corresponding joint inversion models (Figures 4e and f) are visually very similar to the individual inversion models, but the corresponding cross-gradients functions (Figure 4g) are 2-3 orders of magnitude smaller than for the individual inversion models (Figure 4c). In the joint inversion models it is seen that the seismic data have a rather strong influence on the resulting radar wavespeed model in the upper and lower portions of the inversion domain. This is the result of a poor GPR ray coverage in these regions, since many data were discarded due to refractions at the water table at the top and due to the highly attenuating clay at the bottom. Although the scatter plot for the individual inversion models (Figure 4d) shows a strong correspondence between the seismic and radar wavespeeds, as for the Oyster case study, the scatter plot for the joint inversion models (Figure 4h) is defined by much narrower and better defined correlations.

A useful approach for quantifying improvements in resolution is the point-spread functions (PSF), which we calculate following the approach outlined by Alumbaugh and Newman (2000). A PSF can be interpreted as the spatial averaging filter that relates the true underlying model to the resulting inversion model at a specific location for a linearized solution about the final model. Normalization is important for the joint inversion case, in which the calculated PSFs are normalized with respect to the mean values of the radar and seismic slownesses. Figure 5 displays normalized PSF volumes at a central location (x=2.5 m,

$y=2.5$ m, $z=6$ m). These volumes correspond to isosurfaces for which the PSF is 33% of the largest PSF value as suggested by Alumbaugh and Newman (2000). Individual inversions have similar PSFs for the radar (Figure 5a) and seismic (Figure 5b) inversion models. It is clear that the vertical resolution is much higher than the horizontal resolution. When performing joint inversion, the seismic model is at this location influenced by both the seismic (Figure 5c) and radar (Figure 5d) properties in the surroundings. Figures 5e and 5f show the corresponding regions that influence the radar model at this location. These figures demonstrate that the joint inversion may markedly improve the resolution and that the estimated model parameters (e.g., radar or seismic wavespeed) at a given point depend on both the seismic and radar wavespeed fields in the vicinity of this point. The relative resolution improvements by joint inversion are very similar for other choices of isosurfaces (e.g., 15%). Similar results were presented by Linde et al. (2008) for the Oyster case study.

Previous publications on joint inversions based on the cross-gradients function have employed least-squares formulations for data and model misfits. Robust inversions based on l_1 norms are appealing for applications in which the data are noisy or the geology is dominated by a few distinct boundaries, but the computational effort associated with linear programming is often prohibitive. We have investigated the perturbed Ekblom l_p norm for the model norm using IRLS with the common choice of $p=1.0$ and $\gamma=0.1$ (see Equation 6). Figure 6 displays vertical profiles of the seismic and radar wavespeeds obtained from the individual and joint inversions using IRLS and least-squares formulations at $x=5$ and $y=2.5$ m (see Figure 4). The largest differences appear between the joint and the individual inversions, with the joint inversion models showing somewhat more variability. The differences in radar wavespeeds between the individual and joint inversions in the upper part of the model is due to low radar ray coverage (data affected by refractions at the water table were discarded). There are relatively small differences between the joint inversion results obtained using the

IRLS and least-squares model norms; the IRLS inversion results are overall more variable and less smooth. These results illustrate that the joint inversion, at least for the example considered here, have a larger impact on the final inversion results than those related to the model norm used in the inversion.

Seismic and radar wavespeed cross-property relations

The Hashin-Shtrikman (HS) bounds (Hashin and Shtrikman, 1962, 1963) offer an appealing framework for investigating possible relationships between seismic and radar wavespeeds. Without imposing restrictive assumptions about the pore-space geometry, the HS bounds provide the tightest range of the property values a mixture of a two-phase media can take with known volume fractions ϕ and 1-ϕ and properties of each phase. The lower bounds for seismic and radar wavespeed correspond to the case in which spherical inclusions (representing the grains) are embedded in a matrix of water and the spheres are not in contact with each other. The upper bounds correspond to the case where unconnected spherical inclusions of water are imbedded in a solid matrix. Pride et al. (2004) argue that the lower HS bound is generally the one that is the closest to reality in sedimentary settings. Absence of a percolation threshold in most porous media, which indicates that the pore-space is connected down to very low porosities (Sen et al., 1981), supports this argument.

The lower k_{HSL} and upper k_{HSU} HS bounds for bulk modulus in water-saturated media are given by (Hashin and Shtrikman, 1963)

$$k_{\text{HSL}} = k_w + \frac{1-\phi}{\dfrac{1}{k_s - k_w} + \dfrac{3\phi}{3k_w + 4\mu_w}}, \tag{12}$$

$$k_{\mathrm{HSU}} = k_s + \cfrac{\phi}{\cfrac{1}{k_w - k_s} + \cfrac{3(1-\phi)}{3k_s + 4\mu_s}}, \tag{13}$$

where ϕ is porosity, k_s and μ_s are the bulk and shear modulus of the solid, and k_w and μ_w are the corresponding values for the water phase. The lower μ_{HSL} and upper μ_{HSU} bounds for shear modulus are (Hashin and Shtrikman, 1963)

$$\mu_{\mathrm{HSL}} = \mu_w + \cfrac{1-\phi}{\cfrac{1}{\mu_s - \mu_w} + \cfrac{6}{5\mu_w}\cfrac{(k_w + 2\mu_w)\phi}{(3k_w + 4\mu_w)}}, \tag{14}$$

$$\mu_{\mathrm{HSU}} = \mu_s + \cfrac{\phi}{\cfrac{1}{\mu_f - \mu_s} + \cfrac{6}{5\mu_s}\cfrac{(k_s + 2\mu_s)(1-\phi)}{(3k_s + 4\mu_s)}}. \tag{15}$$

The lower and upper bounds for P-wavespeed are then given by ($\mu_w = 0$)

$$\alpha_{\mathrm{HSL}} = \sqrt{\frac{k_{\mathrm{HSL}}}{\rho}}, \tag{16}$$

$$\alpha_{\mathrm{HSU}} = \sqrt{\frac{1}{\rho}\left(k_{\mathrm{HSU}} + \frac{4}{3}\mu_{\mathrm{HSU}}\right)}. \tag{17}$$

The lower κ_{HSL} and upper κ_{HSU} bounds for dielectric permittivity in saturated media are (e.g., Hashin and Shtrikman, 1962; Brovelli and Cassiani, 2009)

$$\kappa_{\mathrm{HSL}} = \kappa_s + \frac{\phi}{\dfrac{1}{\kappa_w - \kappa_s} + \dfrac{1-\phi}{3\kappa_s}}, \tag{18}$$

$$\kappa_{\mathrm{HSU}} = \kappa_w + \frac{1-\phi}{\dfrac{1}{\kappa_s - \kappa_w} + \dfrac{\phi}{3\kappa_w}}, \tag{19}$$

where κ_s and κ_w are the dielectric permittivities of the solid and the water phases. It is then possible to determine the lower v_{HSL} and upper v_{HSU} bounds for radar wavespeed using

$$v_{\mathrm{HSL}} = \frac{c}{\sqrt{\kappa_{\mathrm{HSU}}}}, \tag{20}$$

$$v_{\mathrm{HSU}} = \frac{c}{\sqrt{\kappa_{\mathrm{HSL}}}}, \tag{21}$$

where $c = 3 \times 10^8$ m/s is the vacuum speed of light.

Figures 7a-b display the HS bounds for the radar and seismic wavespeed, respectively, for the case of varying ϕ with: k_s=38 MPa, k_w=2.09 GPa, μ_s=41.5 GPa, μ_w=0, κ_s=6.5, κ_w=84, ρ_s=2.65 kg·m^{-3} and ρ_w=1000 kg·m^{-3}. These values are representative values of α-quartz (Schön, 1996) and water at 10° C (Eisenberg and Kauzmann, 1969; Fine and Millero, 1973). The corresponding relationships between the radar and seismic wavespeeds as a function of porosity are shown in Figure 7c, together with the Oyster and Thur River scatter plots that result from the joint inversions. The scatter plots lie along or very close to the lower HS bounds. It is well-known that tomograms underestimate the variability of the real physical fields (e.g., Day-Lewis and Lane, 2004). Because the estimated radar and seismic slownesses

are based on the same inversion processes, they are approximately equally affected by this limitation. As a consequence, we expect the cross-property center points and slopes (as revealed by the scatter plots) to be more robust descriptions of the system than the tomograms themselves. For the two case studies, we conclude that the pore space is well-connected at both locations and that the Oyster site has significantly higher porosities than the Thur River site.

An example of the averaging that takes place during inversion is demonstrated in Figure 7d in which two types of estimates of porosity variations are shown. One is based on Neutron-Neutron (NN) logs recorded in a borehole located at the center of the Thur River inversion domain (x=2.5 m, y=2.5 m) and one is based on the individual and joint inversion wavespeed models at the same location. The NN-to-porosity transform was obtained following Barrash and Clemo (2002), where the lowest and highest NN counts out of 18 borehole logs at the Thur River site are assigned to the highest (0.50) and lowest (0.12) expected end-member porosities for this type of sedimentary setting. Although the resulting absolute porosities obtained from this type of transform might be biased and the variability over-estimated, the relative variations with depth are expected to be well-resolved. A site-specific NN-to-porosity transform obtained by measuring the porosities on retrieved cores would have helped to improve the absolute porosity values, but no undisturbed cores could be retrieved at our site. The co-located radar and seismic wavespeeds were transformed to porosity via the lower Hashin-Shtrikman bounds using the same parameters as assumed in the construction of Figure 7a-b. Note, that the choice of κ_s=6.5 was treated as a fitting parameter to assure consistent porosity estimates from the seismic and radar joint inversion models. It is seen that the wavespeed models provide plausible and fairly tight lower bounds of porosity, but that only the main trends of porosity as defined by the NN-logs are resolved. A much better correspondence between the overall NN-derived porosity values and those obtained from the

radar wavespeed model is obtained using the volume-averaging approach of Pride (1994) (Klotzche et al., submitted). A possible approach to improve the models would be to include the porosity estimates defined by the NN-logs into the reference model (e.g., Yeh et al., 2002) or to perform full waveform inversion (Klotzsche et al., submitted). Correlation of the radar and seismic wavespeeds with the NN-determined porosities at this location is slightly improved by the joint inversion (the correlation coefficient is increased with 10-15%) and the consistency between the two estimates obtained from joint inversion makes it easier to interpret the results. That the scatter plots appear on the lower HS bounds indicate that the effective porosity is rather similar to the total porosity estimated from the NN logs.

DISCUSSION

The structural approach to joint inversion using cross-gradients constraints (Gallardo and Meju, 2003; 2004) is a maturing inversion technique that might provide internally consistent geophysical models with improved resolution compared with those obtained from individual inversions. The results presented here and elsewhere indicate that joint inversion using cross-gradients constraints may improve: (1) zonation of lithological sub-units (Gallardo and Meju, 2003; 2004; 2007; Doetsch et al., submitted); (2) ratios of physical properties (Tryggvason and Linde, 2006); (3) petrophysical inferences (Linde et al., 2006); (4) field-scale correlations with hydrological properties (Linde et al., 2008). These results suggest that joint inversion based on cross-gradients constraints might one day become a standard tool in diverse multi-method geophysical applications. There are nevertheless several questions that merits further attention and they are briefly discussed below.

How to justify the assumption of structural similarity for a given field application? Some knowledge about the field site is very important. An example of when the joint inversion approach is invalid would be in a heterogeneous geological media with strong

gradients in state variables (e.g., salinity) such as in a coastal setting or at a contaminated site as discussed in Linde et al. (2006). Access to geophysical logging data makes it possible to investigate structural similarity in the vertical direction at a few positions. Note that structural similarity is imposed at the resolution of the resulting models, not at the typically higher resolution of the logging data. We recommend to always performing both individual and joint inversion of given data sets. If the joint inversion fits the data to the same level as the individual inversions, if the scatter plots of the resulting models display the same main trends, and if the joint inversion models appear more distinct but that no fundamentally new structure is added, then joint inversion with cross-gradients constraints might be a valid approach. If not, it might be possible to impose structural similarity in parts of the model or to decrease the weight given to this constraint. In some cases, it might be possible to reformulate the inverse problem using, for example, time-lapse data to better constrain the properties that are expected to vary. There are many conditions when the assumption of structural similarity of model parameters is invalid and a careful analysis is needed for each new application.

What is the best discretization of the cross-gradients function? Instead of discretizing using the neighboring model cells, it might be better to define a discretization on the same scale as for the model regularization. This might further stabilize the joint inverse problem and decrease the sensitivity with respect to the model discretization.

How to determine optimal weights of the components associated with each data set in the objective function and what are the associated trade-offs? Our approach consists of first giving equal weights to each data set and then reweight until all data types can be fitted to the same error level as for the individual inversions. There are many alternative approaches, in which one could consider the resolution properties, spatial coverage, etc. It would probably be quite instructive to perform a detailed analysis of the trade-offs associated with different data types and the different components of the objective function.

How to transfer the joint inversion results to geological models and how to make robust petrophysical inferences? Gallardo and Meju (2003) suggested a manual lithological classification guided by the scatter plots of the different models. Doetsch et al. (2008) developed a formalized classification scheme and performed a zonal inversion for effective petrophysical properties of each zone. A key step in all interpretations is to know what field-scale petrophysical relationships should be used to transform the models into geological or hydrogeological properties. It is useful to employ petrophysical relationships that share similar parameterizations and assumptions about the pore structure for all model types as done here for the HS lower bounds, using a weighted average of the HS bounds (Brovelli and Cassiani, 2010), or by using volume-averaging (Linde et al., 2006). Even if the joint inversion improves resolution, similar resolution-dependent petrophysical relationships as for individual inversions remain. A possibility is to focus on the slope of the cross-property relations that might be better resolved than the variability of each model. It would also be useful to extend the method of Day-Lewis et al. (2005) to this type of problems and thereby improve the determination of field-scale petrophysical relationships. One common assumption in petrophysical inference is that some properties like the dielectric permittivity of the solid phase is constant. One possibility is to relax this assumption through Monte Carlo simulations as suggested by Linde et al. (2006). This might help to better understand different possible explanations, for example, of the different slopes in the low wavespeed region of Figure 4h.

CONCLUSIONS

Joint inversion of crosshole radar and seismic traveltimes based on cross-gradients constraints using least-squares or l_1-norm mimicking measures offers a reliable and robust methodology for improving model resolution in saturated unconsolidated media. With better resolved models, the confidence in subsequent geophysical-petrophysical analyses is

increased. This inversion approach is also expected to be applicable to consolidated sedimentary environments, because porosity is the controlling factor for both radar and seismic wavespeed and the two inversion properties are expected to have similar responses to changes in the pore structure. For consolidated media, it might be useful to define a 1D reference model and to solve for the model update. No applications of joint inversion of radar and seismic data have been reported under multi-phase conditions. Such applications hold considerable promise in the vadose zone and in petroleum exploration applications, but a detailed assessment of the validity of the cross-gradients function under such conditions remains to be investigated. Similar arguments apply to surface-based data.

It is relatively straightforward to extend our joint inversion scheme to include improved forward modeling algorithms based on full-waveform or Fresnel-zone modeling approaches. Significant improvements in hydrogeophysical characterization are usually obtained by also including information about the electrical conductivity distribution obtained by incorporating radar attenuation, full-waveform modeling of radar data, or geoelectrical datain the joint inversion.

REFERENCES

Ajo-Franklin, J. B., B. J. Minsley, and T. M. Daley, 2007, Applying compactness constraints to differential traveltime tomography: Geophysics, **72**, R67-R75.

Alumbaugh, D. L., and G. A. Newman, 2000, Image appraisal for 2-D and 3-D electromagnetic inversion: Geophysics, **65**, 1455-1467.

Barrash, W., and T. Clemo, 2002, Hierarchical geostatistics and multifacies systems: Boise Hydrogeophysical Research Site, Boise, Idaho: Water Resources Research, **38**, 1196.

Belina, F. A., J. R. Ernst, and K. Holliger, 2009, Inversion of crosshole seismic data in heterogeneous environments: Comparison of waveform and ray-based approaches: Journal of Applied Geophysics, **68**, 85-94.

Bradford, J. H., W. P. Clement, and W. Barrash, 2009, Estimating porosity with ground-penetrating radar reflection tomography: A controlled 3-D experiment at the Boise Hydrogeophysical Research Site: Water Resources Research, **45**, W00D26.

Brovelli, A., and G. Cassiani, 2010, A combination of the Hashin-Shtrikman bounds aimed at modelling electrical conductivity and permittivity of variably saturated porous media: Geophysical Journal International, **180**, 225-237.

Chen, J. S., S. Hubbard, Y. Rubin, Y. Murray, C. Roden, and E. Majer, 2004, Geochemical characterization using geophysical data and Markov Chain Monte Carlo methods: A case study at the South Oyster bacterial transport site in Virginia: Water Resources Research, **40**, W12412.

Day-Lewis, F. D., and J. W. Lane, Jr., 2004, Assessing the resolution-dependent utility of tomograms for geostatistics: Geophysical Research Letters, **31**, L07503.

Day-Lewis, F. D., K. Singha, and A. M. Binley, 2005, Applying petrophysical models to radar travel time and electrical resistivity tomograms: Resolution-dependent limitations: Journal of Geophysical Research, **110**, B08206.

Deutsch, C. V., and A. G. Journel, 1998, GSLIB: Geostatistical Software Library and User's Guide: Oxford Univ. Press.

Diem, S., T. Vogt, and E. Hoehn, Räumliche kleinmassstäbliche Verteilung der hydraulischen Leitfähigkeit in voralpinen alluvialen Schotter-Grundwasserleitern: Submitted to Grundwasser.

Dietrich, C. R., and G. N. Newsam, 1997, Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix: SIAM Journal of Scientific Computing, **18**, 1088-1107.

Doetsch, J., I. Coscia, S. Greenhalgh, N. Linde, A. G. Green, and T. Günther, 2010, The borehole-fluid effect in electrical resistivity imaging: Accepted for publication in Geophysics.

Doetsch, J., N. Linde, I. Coscia, S. Greenhalgh, and A. G. Green, submitted, Zonation for 3D aquifer characterization based on joint inversions of multi-method crosshole geophysical data: Geophysics, GEO-2009-0424.

Eisenberg, D., and W. Kauzmann, 1969, The structure and properties of water: Oxford university press.

Ernst, J. R., A. G. Green, H. Maurer, K. Holliger, 2007, Application of a new 2D time-domain full-waveform inversion scheme to crosshole radar data: Geophysics, **72**, J53-J64.

Farquharson, C. G., 2008, Constructing piecewise-constant models in multidimensional minimum-structure inversions: Geophysics, **73**, K1-K9.

Fine, R. A., and F. J. Millero, 1973, Compressibility of water as a function of temperature and pressure: The Journal of Chemical Physics, **59**, 5529-5536.

Fregoso, E., and L. A. Gallardo, 2009, Cross-gradients joint 3D inversion with applications to gravity and magnetic data: Geophysics, **74**, L31-L42.

Gallardo, L. A., 2007, Multiple cross-gradient joint inversion for geospectral imaging: Geophysical Research Letters, **34**, L19301.

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**, 1658.

Gallardo, L. A., and M. A. Meju, 2004, Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints: Journal of Geophysical Research-Solid Earth, **109**, B03311.

Gallardo, L. A., and M. A. Meju, 2007, Joint two-dimensional cross-gradient imaging of magnetotelluric and seismic traveltime data for structural and lithological classification: Geophysical Journal International, **169**, 1261-1272.

Gallardo, L. A., M. A. Meju, and M. A. Perez-Flores, 2005, A quadratic programming approach for joint image reconstruction: Mathematical and geophysical examples: Inverse Problems, **21**, 435-452.

Golub, G. H., and C. F. Van Loan, 1996, Matrix computations: John Hopkins University Press.

Haber, E., and D. Oldenburg, 1997, Joint inversion: A structural approach: Inverse Problems, **13**, 63-77.

Hashin, Z., and S. Shtrikman, 1962, A variational approach to the theory of the effective magnetic permeability of multiphase materials: Journal of Applied Physics, **33**, 3125-3131.

Hashin, Z., and S. Shtrikman, 1963, A variational approach to the theory of the elastic behavior of multiphase materials: Journal of the Mechanics and Physics of Solids, **11**, 127-140.

Holliger, K., M. Musil, and H. R. Maurer, 2001, Ray-based amplitude tomography for crosshole georadar data: a numerical assessment: Journal of Applied Geophysics, **47**, 285-298.

Hu, W., A. Abubakar, and T. M. Habashy, 2009, Joint electromagnetic and seismic inversion using structural constraints: Geophysics, **74**, R199-R109.

Hubbard, S., J. Chen, J. Peterson, E. Majer, K. Williams, D. Swift, B. Mailliox, and Y. Rubin, 2001, Hydrogeological characterization of the DOE Bacterial Transport Site in Oyster, Virginia, using geophysical data: Water Resources Research, **37**, 2431-2456.

Hyndman, D. W., and J. M. Harris, 1996, Traveltime inversion for the geometry of aquifer lithologies: Geophysics, **61**, 1728-1737.

Klotzsche, A., J. van der Kruk, G. Meles, J. Doetsch, H. Maurer, N. Linde, submitted, Full-waveform inversion of crosshole ground penetrating radar data to characterize a gravel aquifer close to the River Thur, Switzerland: Near Surface Geophysics.

Lanz, E., H. Maurer, A. G. Green, 1998, Refraction tomography over a buried waste disposal site: Geophysics, **63**, 1414-1433.

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 traveltime data: Water Resources Research, **42**, W12404.

Linde, N., A. Tryggvason, J. E. Peterson, and S. S. Hubbard, 2008, Joint inversion of crosshole radar and seismic traveltimes acquired at the South Oyster Bacterial Transport Site: Geophysics, **73**, G29-G37.

Lines, L. R.., A. K. Schultz, and S. Treitel, 1988, Cooperative inversion of geophysical data: Geophysics, **53**, 8-20.

Kowalsky, M. B., S. Finsterle, J. Peterson, S. Hubbard, Y. Rubin, E. Majer, A. Ward, and G. Gee, 2005, Estimation of field-scale soil hydraulic parameters and dielectric parameters through joint inversion of GPR and hydrological data: Water Resources Research, **41**, W11425.

Maurer, H., K. Holliger, and D. E. Boerner, 1998, Stochastic regularization: Smoothness or similarity?: Geophysical Research Letters, **25**, 2889-2892.

Monteiro Santos, F. A., S. A. Sultan, P. Represesas, and A. L. El Sorady, 2006, Joint inversion of gravity and geoelectrical data for groundwater and structural investigation: application to the northwestern part of Sinai, Egypt: Geophysical Journal International, **165**, 705-718.

Musil, M., H. R. Maurer, and A. G. Green, 2003, Discrete tomography and joint inversion for loosely connected or unconnected physical properties: application to crosshole seismic and georadar data sets: Geophysical Journal International, **153**, 389-402.

Paige, C. C., and M. A. Saunders, 1982, LSQR: An algorithm for sparse linear equations and sparse least squares: ACM Transactions on Mathematical Software, **8**, 43-71.

Paasche, H., and J. Tronicke, 2007, Cooperative inversion of 2D geophysical data sets: A zonal approach based on fuzzy c-means cluster analysis: Geophysics, **72**, A35-A39.

Paasche, H., A. Wendrich, J. Tronicke, and C. Trela, 2008, Detecting voids in masonry by cooperatively inverting P-wave and georadar traveltimes: Journal of Geophysics and Engineering, **5**, 256-267.

Paige, C. C., and M. A. Saunders, 1982, LSQR: an algorithm for sparse linear equations and sparse least squares: Transactions of Mathematical Software, **8**, 43-71.

Podvin, P., and I. Lecomte, 1991, Finite difference computation of traveltimes in very contrasted velocity models: a massively parallel approach and its associated tools: Geophysical Journal International, **105**, 271-284.

Pratt, R. G., 1999, Seismic waveform inversion in the frequency domain, Part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888-901.

Pride, S., 1994, Governing equations for the coupled electromagnetics and acoustics of porous media: Physical Review B, **50**, 15678-15696.

Pride, S. R., J. G. Berryman, and J. M. Harris, 2004, Seismic attenuation due to wave-induced flow: Journal of Geophysical Research, **109**, B01201.

Ranguelova, E. B., 2002, Segmentation of textured images on three-dimensional lattices: Ph.D. thesis, University of Dublin.

Sen, P. N., C. Scala, and M. H. Cohen, 1981, A self-similar model for sedimentary rocks with application to the dielectric constant of fused glass beads: Geophysics, **46**, 781-795.

Schön, J. H., 1996, Physical properties of rocks: Fundamentals and principles of petrophysics: Elsevier.

Tryggvason, T., and B. Bergman, 2006, A traveltime reciprocity inaccuracy in the *time3d* finite difference algorithm by Podvin & Lecomte: Geophysical Journal International, **165**, 432-435.

Tryggvason, A., and N. Linde, 2006, Local earthquake (LE) tomography with joint inversion for P- and S-wave velocities using structural constraints: Geophysical Research Letters, **33**, L07303.

Tryggvason, A., S. T. Rögnvaldsson, and Ó. G. Flóvenz, 2002, Three-dimensional imaging of the P- and S-wave velocity structure and earthquake locations beneath Southwest Iceland: Geophysical Journal International, **151**, 848-866.

Vasco, D. W., J. E. Peterson, and E. L. Majer, 1995, Beyond ray tomography – wavepaths and Fresnel volumes: Geophysics, **60**, 1790-1804.

Vozoff, K., and D. L. P. Jupp, 1975, Joint inversion of geophysical data: Geophysical Journal of the Royal Astronomical Society, **42**, 977-991.

Yeh, T. C. J., S. Liu, R. J. Glass, K. Baker, J. R. Brainard, D. Alumbaugh, and D. LaBrecque, 2002, A geostatistically based inverse model for electrical resistivity surveys and its applications to vadose zone hydrology: Water Resources Research, **38**, 1278.

Zhdanov, M. S., 2009, New advances in regularized inversion of gravity and electromagnetic data: Geophysical Prospecting, **57**, 463-478.

APPENDIX A: Stochastic regularization operators

The model covariance matrix $\mathbf{C}_m$ is a symmetrical Toeplitz matrix if the correlation function is stationary and the grid discretization is uniform in each direction (e.g., Dietrich and Newsam, 1997). Linde et al. (2006a) used circulant embedding and the diagonalization theorem for circulant matrices to compute the stochastic regularization operator $\mathbf{C}_m^{-0.5}$ as described below. Their method for calculating $\mathbf{C}_m^{-0.5}$ is computationally efficient, because operations are performed on a vector instead of on a matrix.

Following Dietrich and Newsam (1997), $\mathbf{C}_m$ of a stationary process $Y(x)$ with correlation function $C(x)$ sampled on a uniform 1D mesh $\Omega=\{x_0,\ldots,x_m\}$ has values $R_{qr}=r(|x_0-x_k|)$. A model covariance matrix $\mathbf{C}_m$ of size $m \times m$ can be circulantly embedded into a symmetric circulant matrix $\mathbf{S}$ of size $2M \times 2M$ by assigning the following entries to the first column $\mathbf{s}$ of $\mathbf{S}$

$$
\begin{aligned}
s_k &= r_k, & k &= 0,\ldots,m, \\
s_{2M-k} &= r_k, & k &= 1,\ldots,m-1,
\end{aligned}
\tag{A1}
$$

where if $M>m$ the values $s_{m+1},\ldots,s_{2M-m}$ are arbitrary or conveniently chosen. The next column of $\mathbf{S}$ can be obtained by shifting the first column circularly, such that the last element becomes first and all other elements are shifted forward by one, and so on. Being circulant $\mathbf{S}$ can be decomposed by using the diagonalization theorem of circulant matrices

$$
\mathbf{S} = \frac{1}{2M}\mathbf{F}\mathbf{\Lambda}\mathbf{F}^{\mathrm{H}},
\tag{A2}
$$

where $\mathbf{F}$ is the fast Fourier transform (FFT) matrix of size $2M$ with entries $F_{pq}=\exp(2\pi iqr/2M)$, $\mathbf{F}^{\mathrm{H}}$ is the conjugate transpose of $\mathbf{F}$, and $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal entries form the vector $\tilde{\mathbf{s}}=\mathbf{Fs}$ (e.g., Golub and van Loan, 1996). The matrix $\mathbf{S}$ is nonnegative definite if all entries of $\tilde{\mathbf{s}}$ are nonnegative. These results are extendable to two- and three dimensions (Ranguelova, 2002).

The matrix $\mathbf{S}^{-1/2}$ is for the one-dimensional case also circulant and its first column can be obtained as $\mathbf{F}^H \tilde{\mathbf{s}}^{-1/2}$ and the values corresponding to the first column of $\mathbf{C}_m^{-0.5}$ can be retrieved from entries 1 to m (see Equation A1), whereas all other columns of $\mathbf{C}_m^{-0.5}$ can be calculated by shifting the first column circularly. To decrease memory requirements, only elements of $\mathbf{C}_m^{-0.5}$ that are larger than 1% of the maximum value of $\mathbf{C}_m^{-0.5}$ are stored. In three dimensions, the only difference is that $\mathbf{s}$ and $\tilde{\mathbf{s}}$ are expressed as three-dimensional arrays and that three-dimensional FFT is applied.

To ensure that $\mathbf{S}$ is nonnegative definite in three dimensions when using an exponential correlation function, it is necessary to choose M to be at least seven integral scales in each direction and choose $s_{m+1},...,s_{2M-m}$ to be the corresponding values of $r(l)$. The exponential covariance model used to calculate the entries R_{qr} above is for a stationary three-dimensional domain defined as

$$C(l) = c\mathrm{e}^{-l}, \tag{A3}$$

where c is the variance, e is the natural logarithm, and l is defined as

$$l = \sqrt{\left(\frac{h_x}{I_x}\right)^2 + \left(\frac{h_y}{I_y}\right)^2 + \left(\frac{h_z}{I_z}\right)^2}, \tag{A4}$$

where h_x (m), h_y (m), and h_z (m) are the lags (i.e., the distances between a given pair of model parameters) in the x-, y- and z-direction, respectively, and I_x (m), I_y (m), and I_z (m) are the integral scales specified in the text for the different examples (i.e., the distance at which the correlation between model parameters is 1/e) in the x-, y- and z-direction, respectively.
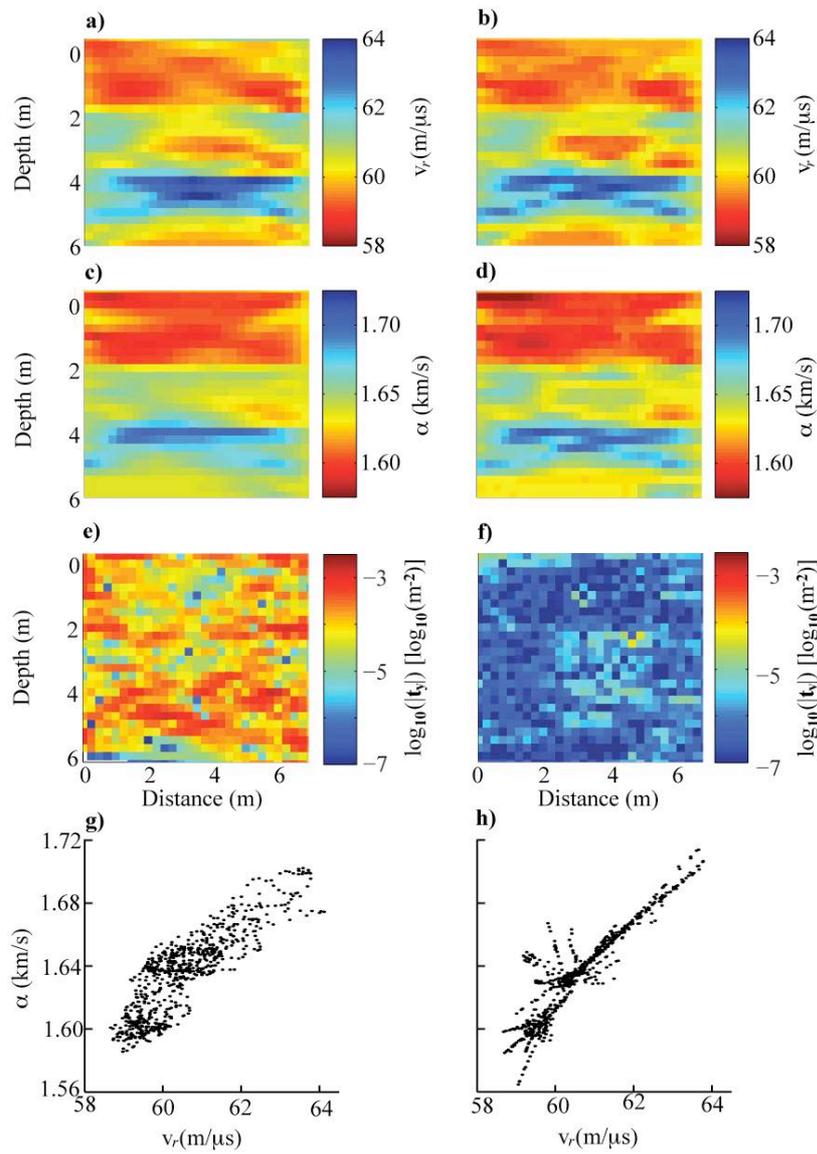
Figure 1: Radar wavespeed models from the Oyster site: (a) individual inversion with stochastic regularization; (b) joint inversion with stochastic regularization; (c) and (d) corresponding seismic wavespeed inversion results; (e) and (f) cross-gradients functions for these models; (g) and (h) scatter plots for these models. Depths are given in meters below sea level. Modified from Figure 3 in Linde et al. (2008).
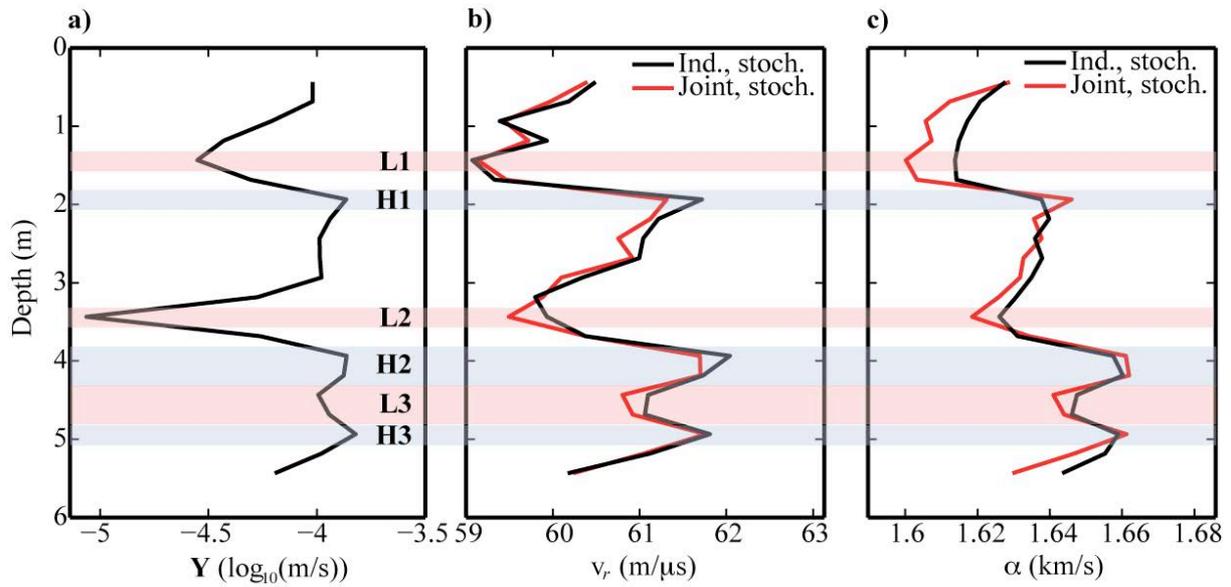
Figure 2: (a) Hydraulic conductivity data from borehole M3 at the Oyster site (located on the right side of the tomogram in Figure 1); (b) tomographic radar wavespeed models located two model cells from M3; (c) tomographic seismic wavespeed models located two model cells from M3. The black and red solid lines in (b) and (c) represent models from the individual and joint inversion models with stochastic regularization. The shaded zones (L1-L3) and (H1-H3) are locations at which hydraulic conductivities have local minima and maxima. Modified from Figure 6 in Linde et al. (2008).
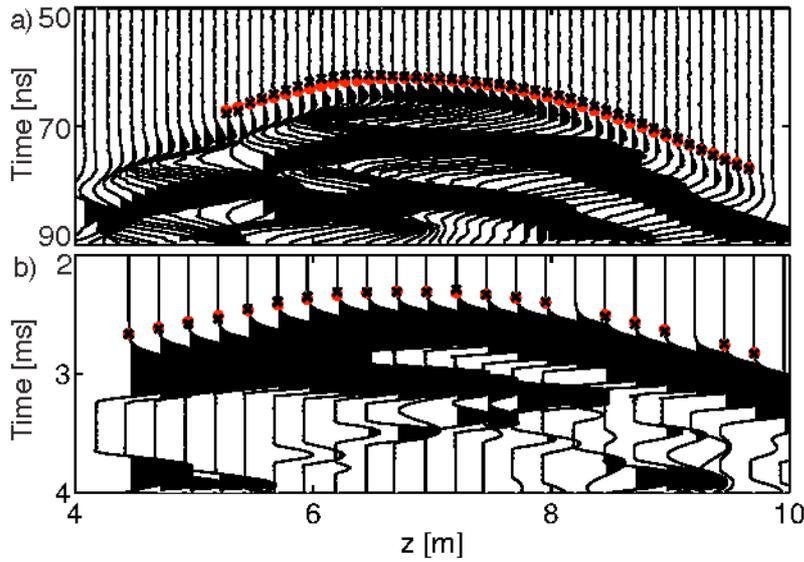
Figure 3: Typical raw (a) seismic and (b) radar source gathers for a source depth of 6.75 m. Red dots in (a) and (b) represent calculated forward responses of the final models obtained by joint inversion (see Figure 4e and 4f) and black crosses represent the picked first arrivals. (a) Although the seismic data were clipped, first arrivals could be reliably picked. (b) Picked first arrivals in the radar data do not include refracted waves through the unsaturated high wavespeed layer above 4 m; for the displayed source gather this means neglecting data collected above 5 m depth.
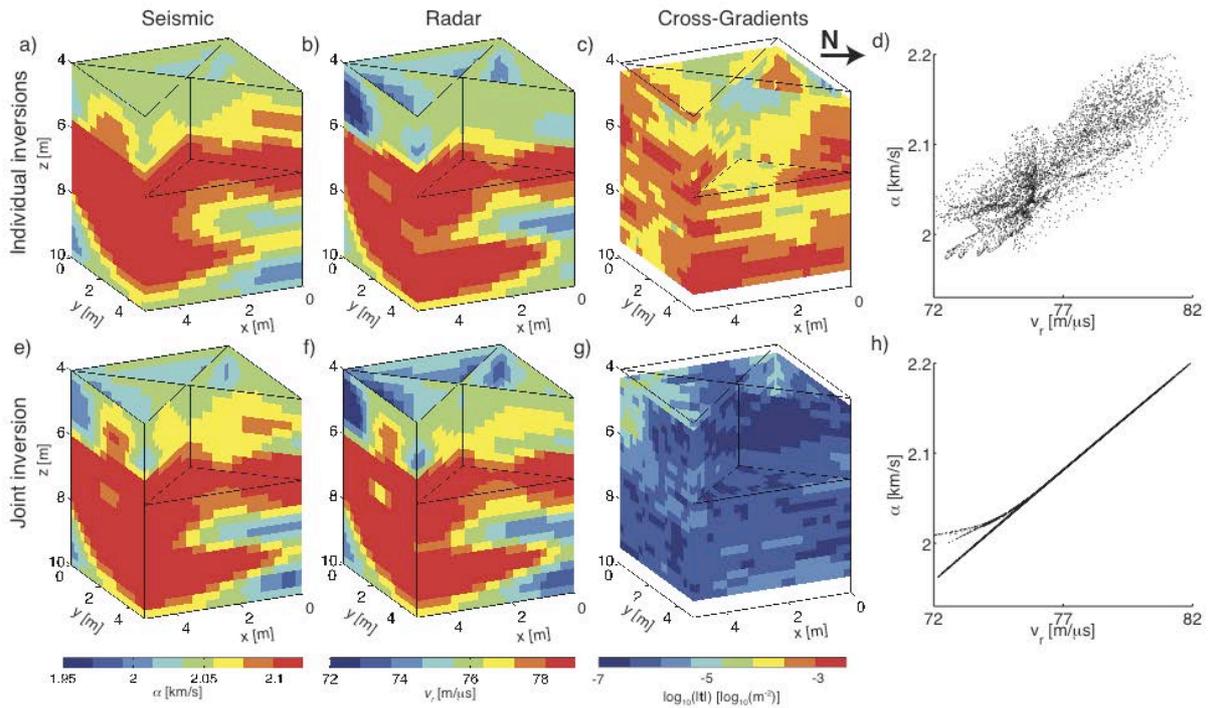
Figure 4: (a) Seismic and (b) radar wavespeed models determined from individual inversions of the Thur River site data; (e) and (f) corresponding models determined from joint inversion; (c) and (g) the cross-gradients functions for these models; (d) and (h) scatter plots for these models.
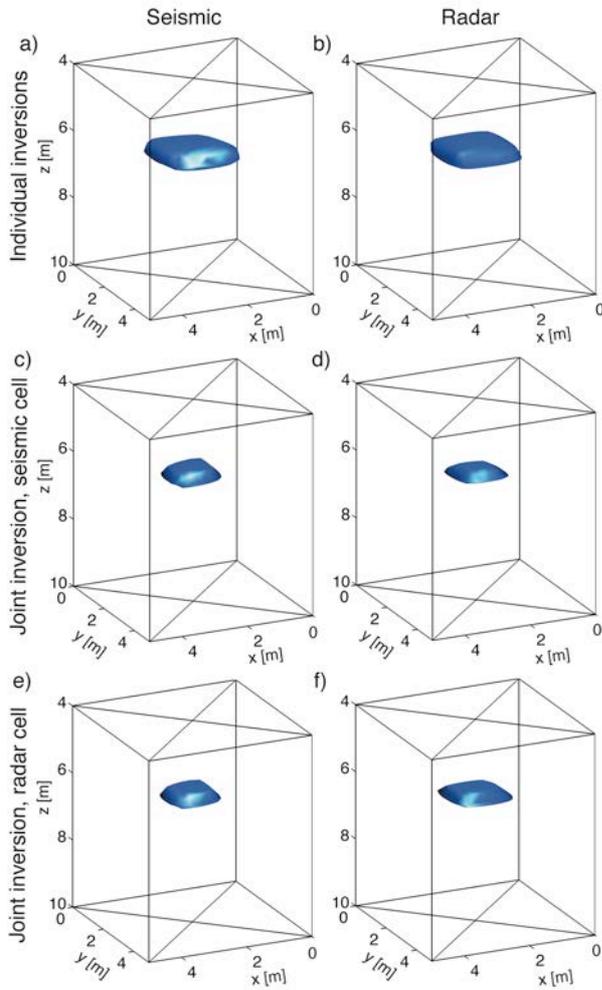
Figure 5: Normalized point-spread functions (PSFs) for the individual (a) seismic and (b) radar inversion models at location x=2.5 m, y=2.5 m, z=6 m for the Thur River site (see Figure 3), where the volume is the region in which the values of the PSFs are at least 33% of the values at the model cell of interest. Normalized PSFs for the seismic model obtained by joint inversion have a smaller spatial support (c), but are also influenced by the radar model (d) over a similar region. Corresponding PSFs for a radar cell showing the influence of the (e) seismic model and the (f) radar model, respectively.
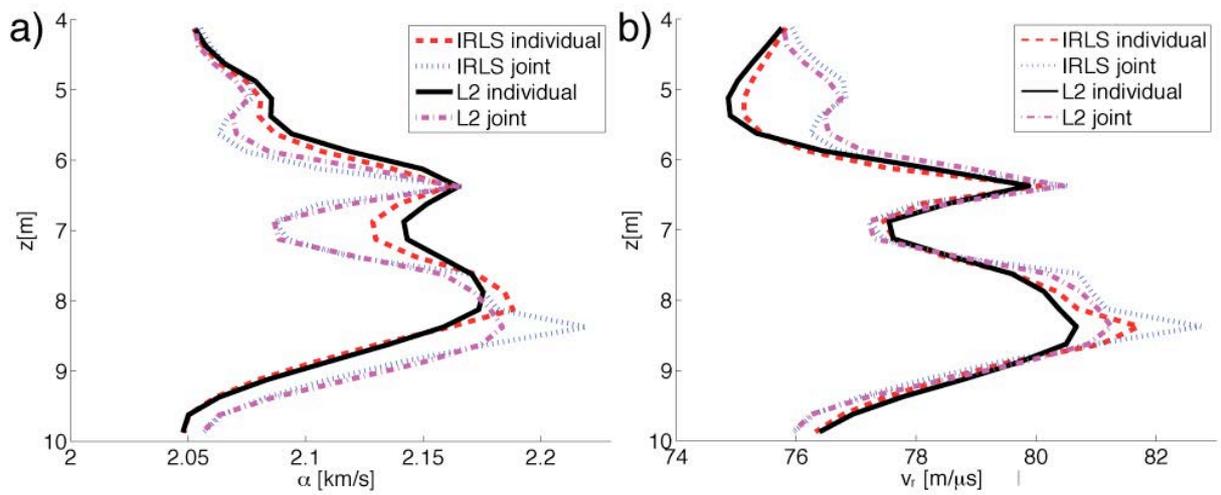
Figure 6: Comparison of (a) seismic and (b) radar wavespeed models at location *x*=5.0 m and *y*=2.5 m (see Figure 3) obtained by individual and joint iteratively reweighted least-squares (IRLS) and least-squares (L2) inversions.
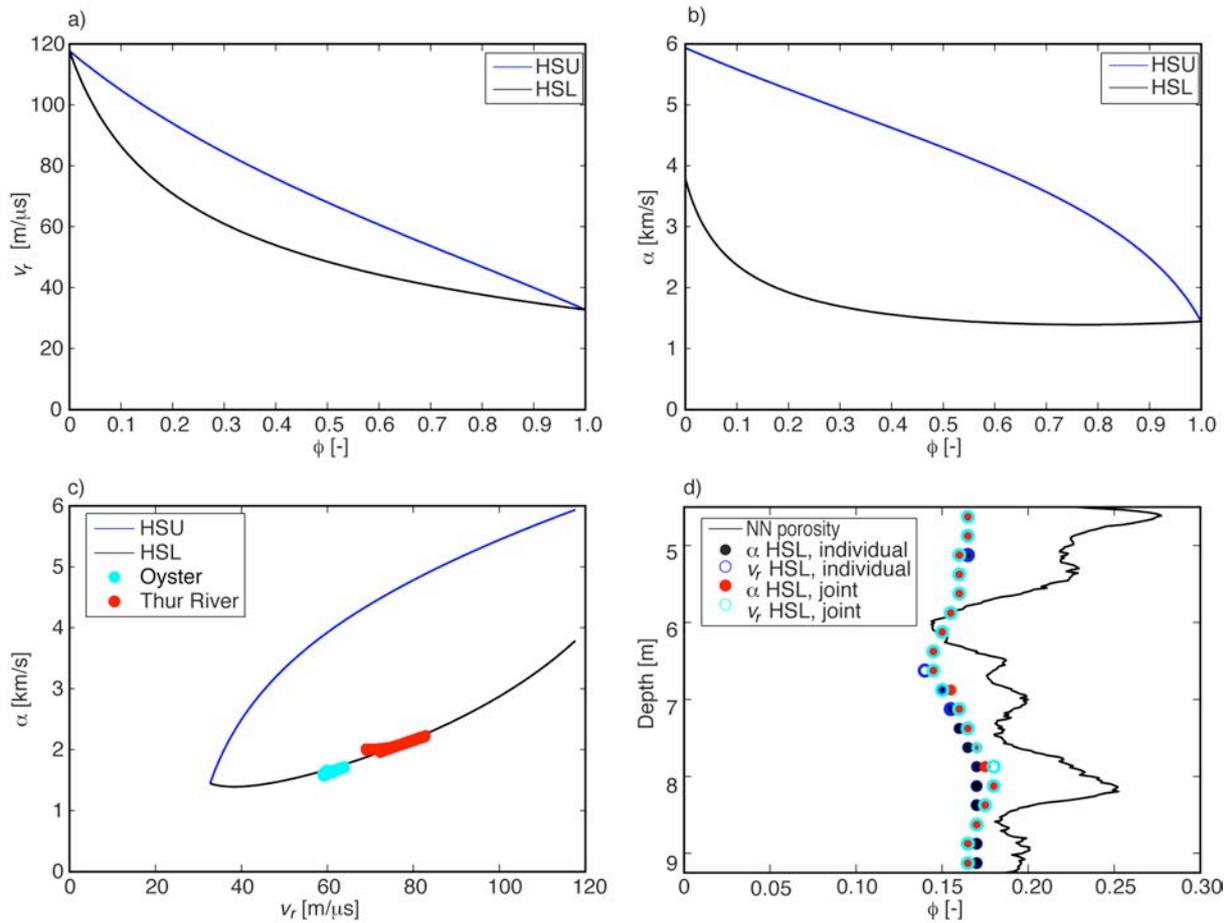
Figure 7: For the parameters described in the text: (a) Hashin-Shtrikman upper (HSU) and lower (HSL) bounds for seismic wavespeed as a function of porosity; (b) HSU and HSL bounds for radar wavespeed; (c) radar-seismic wavespeed relationships for the HSU and HSL bounds as a function of porosity together with scatter plots from the Oyster and Widen case studies; (d) porosity log derived from Neutron-Neutron data and porosity estimates obtained from radar and seismic wavespeed using the HSL bounds for the seismic and radar wavespeeds at the Thur River site.