Joint probabilistic inversion of DC resistivity and seismic refraction data applied to bedrock/regolith interface delineation

G. de Pasquale[1], N. Linde[1] and A. Greenwood[1]

[1]Applied and Environmental Geophysics Group, Institute of Earth Sciences, University of Lausanne

Key Points:

- Probabilistic inversion results better describe bedrock properties
- Joint inversion provides more accurate interface estimates
- Joint inversions have an enhanced sensitivity to model errors

Corresponding author: Niklas Linde, Niklas.linde@unil.ch

Abstract

The depth and geometry of the bedrock/regolith interface influence many near-surface processes. Geophysics offer means to image them non-invasively since they correspond to discontinuities in physical properties. For instance, refraction seismic applications can leverage the strong correlation between seismic velocity and porosity, while DC resistivity measurements with their sensitivity to electrical conductivity are indirectly sensitive to porosity and clay content. Combining these two data types is beneficial since their spatial resolution patterns and sensitivities to relevant target properties differ. Here, we propose a probabilistic joint inversion framework that uses DC resistivity and first-arrival seismic travel times to infer a common interface geometry separating two heterogeneous sub-domains. A synthetic test case demonstrates that the method is more accurate in sampling the target interface than inversion results obtained using single datasets. We then applied the joint inversion method to field data from the Calhoun Critical Zone Observatory in South Carolina, where it resolves apparent contradictions in the individual inversion results. It produces a bedrock topography that mirrors surface topography and offers clear indications of a more fractured bedrock below the surface topography lows.

keywords Inverse theory, joint inversion, Bayesian inversion, bedrock interface, electrical resistivity tomography (ERT), refraction seismics

1 Introduction

The depth to the bedrock and its topography influence many surface and near-surface processes [*Rempe & Dietrich*, 2014] such as water runoff to channels [*Onda et al.*, 2004] and water chemistry [*Anderson et al.*, 2002]. Furthermore, the bedrock/regolith interface is a key boundary when studying erosion rates and certain atmospheric processes related to plant water uptake and transpiration [*Jones & Graham*, 1993]. Its characteristics are crucial for modeling soil thickness, which depends on the balance between production at depth (i.e., where weathering of bedrock takes place) and erosion at the surface [*Heimsathet et al.*, 1997]; an interplay which ultimately determines the critical zone (CZ) thickness [*Brantley et al.*, 2007]. Moreover, this boundary is a key controlling factor for landscape evolution and landslide investigations [*Lanni et al.*, 2013; *Reneau & Dietrich*, 1987]. For these reasons, its imaging is beneficial for a wide range of applications. Geophysical investigations have become a common tool in geomorphological research [*Schrott et al.*,

2003], because they allow for non-invasive and relatively fast surveys that are sensitive to discontinuities in physical properties [*Sass*, 2007]. Moreover, geophysical measurements provide complementary information about subsurface heterogeneities that is complementary to those offered by remote sensing, geomorphological or geological mapping [*Olona et al.*, 2010; *Tye et al.*, 2011; *Hirsch et al.*, 2007].

One geophysical technique that is commonly used to infer interface geometries is the seismic refraction method, which uses the first-arrival travel times of critically refracted body waves to reconstruct P-wave velocity models of the subsurface [*Zhuo*, 2014]. Given the strong correlation between seismic velocity variations and contrasts in density and porosity, the seismic refraction method is well suited to detect and image subsurface interfaces; for instance, the thickness of weathered layers or fractured zones [*Lee & Freitas*, 1990]. Another geophysical method that is widely used in this respect is electrical resistivity tomography (ERT; *Chambers et al.*, 2012; *Hsu et al.*, 2010). ERT provides comparatively high-resolution images of the shallow subsurface (down to depths of tens to a few hundreds of meters) with changes in electrical resistivity being strongly related to changes in water and clay content [*Dahlin*, 1996]. Their difference in spatial resolution and sensitivity to target properties (e.g., water content) has motivated the combined use of seismic refraction and ERT in numerous studies [*Berge et al.*, 2000; *Linder et al.*, 2010; *Hellman et al.*, 2017; *Juhojuntti & Kamm*, 2015]. It is well established that joint inversions aiming at inferring the two physical fields (electrical resistivity and P-wave velocity) can improve resolution and decrease interpretation ambiguity compared to individual inversions [*Gallardo & Meju*, 2004].

In this paper, we focus on the seismic refraction and DC resistivity datasets that were used by *St. Clair et al.* [2015] to infer the geometry of the interface between the critical zone (CZ) and the intact bedrock at the Calhoun Critical Zone Observatory in South Carolina. In their work, the underlying motivation was to better understand the influence of surface topography on bedrock weathering. To do so, *St. Clair et al.* [2015] inverted the two datasets separately before interpreting them qualitatively. Here, we jointly invert the two datasets by using structural constrains that impose a common interface geometry of the two physical property fields [*Kozlovskaza et al.*, 2007; *Linde & Doetsch*, 2016]. To achieve this, we adapt a probabilistic inversion framework for single-method data aiming at inferring an interface separating two heterogeneous sub-domains [*de Pasquale et al.*, 2019] to two data types with model coupling offered by a common interface. We investi-

gate the added value offered by the joint inversion results compared to those obtained by individual inversions. Based on these results, we seek to complement the findings and interpretations by *St. Clair et al.* [2015], in which ERT and seismic refraction surveys were used to investigate the influence of surface topography on bedrock geometry and properties.

2 Methodology

2.1 Geophysical methods and settings

DC resistivity is a geophysical method that is used to image the subsurface soil and rock resistivity by observations made with electrodes at the Earth's surface or in boreholes [*Binley & Kemna*, 2005]. We use the Boundless Electrical Resistivity Tomography (BERT) library [*Rücker et al.*, 2006; *Günther et al.*, 2006] to simulate the DC resistivity forward response. Using BERT, we can compute the 2.5D forward response by relying on a finite element scheme implemented on unstructured meshes [*Si*, 2015]. Unstructured meshes allow for efficient local refinement and are suitable to account for surface topography and internal boundaries. Given that we consider examples with a significant topography, we rely on geometric factors that are calculated numerically by forward modeling.

Seismic refraction tomography relies on first-arrival travel times corresponding to direct and critically-refracted seismic waves to reconstruct seismic P-wave (i.e., compressional wave) velocity models. The time at which these waves arrive at receiver positions (i.e., first-arrival travel times) are the main observations used in seismic refraction surveys. Besides the density of the rock matrix (related to mineralogical composition and texture), seismic velocities are affected by porosity and pore-fluids, as well as confining stress, pressure and temperature [*Schmitt*, 2015]. In order to simulate the first-arrival travel times, we use the physics refraction class of pyGIMLI (Geophysical Inversion and Modelling Library in Python; *Rücker et al.*, 2017). Specifically, the forward operator used is based on the calculation of the fastest paths from sources to receivers along the elements in the unstructured mesh [*Moser*, 1991; *Dijkstra*, 1959]. This algorithm by Dijkstra is very popular, but also imperfect, since the traveled distance is always overestimated given that rays can only move along the discretized elements. It is used herein because of its speed, but we will see later that the related modeling errors have adverse effects on the joint inversion results.

The field-based case-study considers the datasets that were collected at the Calhoun Critical Zone Observatory in South Carolina and interpreted by *St. Clair et al.* [2015]. Here, DC resistivity and seismic refraction surveys were used to investigate the influence of surface topography on bedrock geometry. The DC resistivity survey was acquired using 56 electrodes spaced 5 m apart with a dipole-dipole electrode configuration. The 400 m long profile was obtained by using one roll-along in which 50% of the electrodes stayed in place, which resulted in a total of 84 different electrode positions. When using the original data set, *de Pasquale et al.* [2019] report that a very fine model parameterization is needed to explain the fine-scale details close to the surface. Since our interest is in the deeper CZ structure, we could save computational time, without significantly affecting the results in the area of interest, by only considering every second electrode position and removing configurations with a maximal electrode separation of 30 m. This resulted in 42 electrodes and 645 measurements being used. For the seismic refraction survey, 96 geophones were used with a spacing of 2.5 m and the shots spaced 10 m apart. This configuration of 237.5 m in length was repeated once after moving the set-up and complemented with off-line shots to link the two acquisitions, resulting in 192 geophone positions. We manually picked the first arrival travel times on each trace after processing the data with a zero-phase Butterworth filter (15 Hz 8dB/octave low cut and 150 Hz 24 dB/octave high cut) to improve the signal-to-noise ratio. Particularly, the signal was weak at far offsets and it was particularly noisy at the end of the profile. Travel time curves were plotted for quality control purposes, whereby erroneous picks were ascertained and deleted from the data through comparison with neighboring travel-time curves. This process resulted in 3896 travel times that were kept for inversion. For the ERT dataset, we used a noise description based on an uncorrelated relative Gaussian error of 3.7 %, while for the refraction seismic observations we considered an absolute error of 3 milliseconds and used these noise levels as target data fits for all individual (deterministic and probabilistic) and joint probabilistic inversions.. These choices were made to obtain similar data misfits as for the inversion results presented by *St. Clair et al.* [2015], thereby, facilitating comparisons. All inversion results presented herein, have near-identical data fits.

To demonstrate the proposed inversion methodology, we first consider a synthetic test case in which DC resistivity and seismic refraction surveys are simulated for a test model with the same measurement configurations and surface topography as for the real dataset. The resulting data are noise-contaminated according to the data error assumed for

the field data and subsequently used to jointly invert for the interface between regolith and intact bedrock.

2.2 Probabilistic joint inversion with interface constraints

de Pasquale et al. [2019] presented a probabilistic formulation and solution to the inverse problem of using one geophysical data type to infer the location and geometric shape of an interface separating two heterogeneous sub-domains. In this empirical-Bayes-within-Gibbs algorithm, the interface geometry and physical property field updates are performed alternately within a Markov chain Monte Carlo (MCMC) scheme [*Sambridge & Mosegaard*, 2002]. Furthermore, the interface geometry and the physical properties of the sub-domains are constrained to favor smooth spatial transitions and pre-defined property bounds. The algorithm by *de Pasquale et al.* [2019] is here adapted to joint inversion in which the interface is considered common for two physical property fields. The joint probabilistic inversion is summarized below in four steps for the problem of delineating the CZ, while a more detailed treatment of algorithmic details is provided by [*de Pasquale et al.*, 2019].

(1) **Chain initialization:** the initial model realizations of seismic P-wave velocity ($\mathbf{m}_0^v$) and electrical resistivity ($\mathbf{m}_0^r$) are obtained from the initial property fields of both the CZ ($\mathbf{m}_0^{r\mathrm{CZ}}$, $\mathbf{m}_0^{v\mathrm{CZ}}$) and bedrock ($\mathbf{m}_0^{r\mathrm{B}}$, $\mathbf{m}_0^{v\mathrm{B}}$) domains that are defined over the whole model domain (Figs. 1a and b) and the initial interface ($\mathbf{I}_0$, Fig. 1c). We initialize the physical properties on a 58×10 grid with cell size of 8×8 meters, by randomly drawing values from uncorrelated log-uniform distributions. These fields are linearly interpolated on the unstructured mesh used for the forward computations (Figs. 1a and b). In this way we contain the dimension and geometric-complexity of the inversion parameters. The interface is sampled as a set of connected nodes within the unstructured mesh and determines the field that is to be used for forward simulations (i.e., CZ above and bedrock below the sampled interface as shown in Figs. 1d and e). Note that the ERT forward simulation mesh is expanded to the sides and with depth to avoid boundary effects. Once $\mathbf{m}_0^v$ and $\mathbf{m}_0^r$ are built, we evaluate their likelihoods [*Tarantola*, 2005]: $L(\mathbf{m}_0^v, \mathbf{I}_0|\mathbf{d}^{\mathrm{RS}})$ and $L(\mathbf{m}_0^r, \mathbf{I}_0|\mathbf{d}^{\mathrm{ERT}})$, where $\mathbf{d}^{\mathrm{RS}}$ and $\mathbf{d}^{\mathrm{ERT}}$ are the first-arrival travel times and apparent resistivities, respectively. Following *de Pasquale et al.* [2019], we evaluate corre-

sponding constraint functions that are used to quantify model parameter variability, $c(S_0^{v,r})$, and interface roughness, $c(R_0)$. Their values are used in the inversion to penalize model structure with weights that are inferred using the empirical Bayes approach as described by *Rosas-Carbajal et al.* [2014]. For both constrain types, we rely on a total variation measure.

(2) **Interface update proposals:** after updating the interface, following a principle of minimum change, both the physical parameter fields (seismic velocity and electrical resistivity) are re-mapped accordingly to the proposed interface. The resulting proposal is accepted with probability:

$$\alpha_I = min\left[1, \frac{L(\mathbf{m}^v_{\text{curr}}, \mathbf{I}_{\text{prop}}|\mathbf{d}^{\text{RS}})L(\mathbf{m}^r_{\text{curr}}, \mathbf{I}_{\text{prop}}|\mathbf{d}^{\text{ERT}})c(R_{\text{prop}})}{L(\mathbf{m}^v_{\text{curr}}, \mathbf{I}_{\text{curr}}|\mathbf{d}^{\text{RS}})L(\mathbf{m}^r_{\text{curr}}, \mathbf{I}_{\text{curr}}|\mathbf{d}^{\text{ERT}})c(R_{\text{curr}})}\right], \tag{1}$$

where the subscripts prop and curr refer to the proposed and current state of the Markov chain.

(3) **Physical property update proposals:** both physical property fields are perturbed for each sub-domain using a symmetric Gaussian proposal distribution and mapped according to the interface obtained in step (2). The resulting resistivity and velocity field proposals are accepted or rejected independently from each other, with probabilities:

$$\alpha_{\text{ERT}} = min\left[1, \frac{L(\mathbf{m}^r_{\text{prop}}, \mathbf{I}_{\text{curr}}|\mathbf{d}^{\text{ERT}})c(S^r_{\text{prop}})}{L(\mathbf{m}^r_{\text{curr}}, \mathbf{I}_{\text{curr}}|\mathbf{d}^{\text{ERT}})c(S^r_{\text{curr}})}\right], \tag{2}$$

$$\alpha_{\text{RS}} = min\left[1, \frac{L(\mathbf{m}^v_{\text{prop}}, \mathbf{I}_{\text{curr}}|\mathbf{d}^{\text{RS}})c(S^v_{\text{prop}})}{L(\mathbf{m}^v_{\text{curr}}, \mathbf{I}_{\text{curr}}|\mathbf{d}^{\text{RS}})c(S^v_{\text{curr}})}\right]. \tag{3}$$

The procedure is then repeated starting from step 2.

(4) **Chain finalization**: after “burn-in" (i.e., the number of iterations needed to start sampling proportionally to the posterior distribution; *Meyn & Tweedie* , 1993) and when the stationary part of the chain has reached the desired length, the iteration loop over steps 2 and 3 terminates and the MCMC realizations after burn-in represent our estimate of the posterior distribution of the investigated subsurface.

A schematic flowchart of the algorithm is presented in Fig. 2.

2.3 Petrophysical relationships

In most rocks and soils, electrical current is due to electrolytic conduction of ions. As a consequence, the main factors affecting subsurface resistivity are water content, salinity, porosity and the connectivity of the water phase [*Lesmes & Friedman*, 2005]. Another important factor affecting electrical resistivity is clay type and content through its contribution to enhanced conduction within the electrical double layer at the mineral-water interface. Among many petrophysical relationships, we consider below the relationship proposed by *Linde et al.* [2006], which is based on volume-averaging in the high-salinity limit (the pore contribution to overall conduction is much larger than the surface contribution), to highlight how electrical conductivity (the inverse of electrical resistivity) is affected by properties and state variables of interest:

$$\sigma_{eff} = \frac{1}{F}\left[S_w^n \sigma_w + (F-1)\sigma_s\right]. \tag{4}$$

Here, σ_{eff} (S/m) is the effective conductivity of the matrix-water-air system at the scale of a Representative Elementary Volume (REV), σ_s the surface conductivity, σ_w the pore water conductivity, $F = \phi^{-m}$ the electrical formation factor which depends on the porosity, ϕ, and on the cementation exponent m, S_w is the water saturation and n is the saturation exponent, which is related to the tortuosity of the water phase.

To quantitatively interpret seismic properties in fractured rock it is common to rely on an equivalent medium representation. Here, the multiphase (minerals, water, air) properties and their connectivities are replaced locally (at the scale of a REV) by an upscaled homogeneous medium with the same macroscopic properties [*Liu & Martinez*, 2012]. In this field of rock physics [*Mavko et al.* , 2009], there are many formulations with different underlying assumptions. They range from macroscopic models of heterogeneous rocks that do not account for microscopic heterogeneities (i.e., pore shape, connectivity and fluid distribution) [*Biot*, 1956; *Gassmann*, 1951], to inclusion-based models which upscale microscopic features to describe the macroscopic elastic behavior of the rock [*Xu*, 1998; *Hudson*, 1981; *Chapman*, 2003]. A rough indication of the impact of porosity in saturated rocks is given by Wyllie's mixing law [*Wyllie et al.*, 1956, 1958]:

$$\frac{1}{V_{eff}} = \frac{\phi}{V_f} + \frac{1-\phi}{V_m}, \tag{5}$$

where V_{eff}, V_f and V_m are the P-wave velocities of the saturated rock, the pore fluids and the rock matrix (mineral matrix). All common rock physics models and observations agree that seismic velocities in fractured rocks are lower than in intact rocks, with

the contrast depending on fracture porosity, connectivity and filling material. For instance, the P-wave velocity in water is five times greater than in air, therefore, water saturated porous/fractured rocks presents a higher elastic wave velocity than unsaturated rocks [*Barton*, 2009; *Nur & Simons*, 1969].

3 Results

3.1 Synthetic test case

In the synthetic test case, we consider identical survey layouts and the same surface topography as for the subsequent case-study at the Calhoun CZO. Moreover, the physical properties are assumed to have a log-uniform prior of resistivity in Ωm and velocity in $\frac{\mathrm{m}}{s}$ with the same prior ranges as later used for the Calhoun case-study. The prior ranges for the physical property fields are $\mathbf{m}^r_{CZ} \in [log_{10}(80), log_{10}(10000)]$, $\mathbf{m}^v_{CZ} \in [log_{10}(300), log_{10}(4000)]$ for the CZ sub-domain and $\mathbf{m}^r_b \in [log_{10}(2000), log_{10}(30000)]$, $\mathbf{m}^v_b \in [log_{10}(4000), log_{10}(6000)]$ for the bedrock. Furthermore, we place the true interface at a similar depth range as for the real case (i.e., between the surface and 20 m depth).

Fig. 3(a) represents the electrode positions and the resistivity model used to generate the synthetic ERT datasets, while Fig. 3(c) shows the interpolation (through the nearest neighbor method) of the input resistivity model on the mesh used for the forward computations. Fig. 3(b) displays the shot positions and the P-wave velocity model, while Fig. 3(d) represents its interpolation on the mesh used for the forward computation. For both velocity and resistivity fields, the two heterogeneous sub-domains are obtained by stationary multivariate Gaussian process generations through circulant embedding of the covariance matrix [*Dietrich & Newsam*, 1997]. Fig. 3(e) shows the ERT inversion model result obtained by a l_1-norm mimicking smoothness-constrained deterministic inversion that do not penalize sharp spatial variations stronger than gradual variations, while Fig. 3(f) shows the corresponding deterministic inversion model result for the refraction seismic observations. In Figs. 3(e) and (f) we indicate the interface obtained by using the maximum vertical gradient method on the deterministic inversion results (following the procedure described by *Chambers et al.*, 2012). Moreover, the deterministic inversion model results are depicted with a transparency that is inversely proportional to the cumulative sensitivity. This highlights that the sensitivity of the seismic refraction measurements deteriorates significantly after 400 m. Therefore, for the analysis and interpretation of the inversion re-

sults, we focus only on the results obtained in the first 400 m. Finally, in Figs. 3(g-h) we show a measure of discrepancy between the target model and the inversion results:

$$\begin{aligned} \mathbf{w}_r &= \mathbf{m}^r_{true} - \mathbf{m}^r, \\ \mathbf{w}_v &= \mathbf{m}^v_{true} - \mathbf{m}^v, \end{aligned} \tag{6}$$

where $\mathbf{m}^r_{true}$ and $\mathbf{m}^v_{true}$ are the vectors of the log10 resistivity and velocity values of the true model (Fig. 3c for the resistivity and Fig. 3d for the velocity); while $\mathbf{m}^r$ and $\mathbf{m}^v$ are the vectors of the log10 resistivity and velocity deterministic inversion results, respectively. The mean absolute value of the model discrepancy is 0.28 for the resistivity and 0.12 for the velocity inversion model results.

In both deterministic inversions, the data were fitted to the same error levels as those used in the subsequent probabilistic inversions of the synthetic data. These correspond to inflating the assumed data errors by a factor of 1.1 with respect to the standard deviation used to contaminate the DC resistivity observations, thereby, yielding a standard deviation of 4.1% and by a factor of 1.7 for the seismic refraction observations, thereby, yielding a standard deviation of 5.1 milliseconds. In an attempt to partly avoid committing an "inverse crime", we used a different unstructured mesh when generating the data compared with the one used for the inversions. This implies that the best model that we can recover for the synthetic example is the ones reproduced by Figs. 3c-d and not the true data-generating model in Figs. 3a-b. The factors mentioned above were determined by comparing the residuals between these two types of models. The addition of this model error would not have been needed if the data generating mesh would have been the same as in the subsequent inversion.

To assess the added value of the proposed probabilistic joint inversion algorithm, we also run the empirical-Bayes-within-Gibbs algorithm by *de Pasquale et al.* [2019] considering the DC resistivity and refraction seismic datasets individually. For each case (i.e., two individual and one joint inversion), to sample the posterior space, we run three independent MCMC chains for 5×10^5 iterations with each iteration taking $\approx$ 0.4 seconds when inverting the seismic dataset, $\approx$ 1.7 seconds when inverting the DC resistivity dataset and $\approx$ 2 seconds for the joint inversion. We start all nine MCMC chains with random uncorrelated initial models and use Geweke analysis [Geweke, 1992] to asses the burn-in period of the chain. This method proceeds by testing if the mean of the log-likelihood of the first part (20 % in our implementation) of the supposedly stationary section of the MCMC

chain can be assumed to be the same as the later part (last 50 % in our implementation). By incrementally increasing the initial part of the chain that is discarded from the analysis, it is possible to estimate the burn-in period. The convergence statistics for the different inversion schemes are summarized in Table 1. Considering the conservative choice of using the last 2×10^5 iterations for which the burn-in has reached by all inversions with a considerable margin, the potential scale reduction factor ($\hat{R}$; *Gelman & Rubin*, 1992) is found to be below 1.2 for the majority of the model parameters of the velocity fields for both the bedrock (100 % of the parameters for individual and 98 % for joint inversion) and the CZ sub-domains (69 % for seismic refraction and 61 % for joint inversion). The resistivity field parameters have $\hat{R} < 1.2$ for the majority of the bedrock parameters (70 % for individual and 68 % for joint inversion) but not for the CZ sub-domains (20 % for individual and 24 % for joint inversion). Since this threshold value is typically used to assess if the MCMC chains have sampled the posterior sufficiently, we find that we did not sufficiently sample the posterior pdf, in particular, the CZ resistivity fields. This implies that each individual chain has not sufficiently sampled the posterior pdf. Consequently, the spread of the parameters is likely larger than those inferred in each chain. Nevertheless, combining the three chains for each inversion case increases the standard deviation of the model parameters. For the velocity fields, the parameters of the bedrock sub-domain have a slight augmentation in standard deviation (average of 1% for individual and 7% for joint inversion), while for the CZ sub-domain this increase is more significant (average of 26% for individual and 32% for joint inversion). In general the resistivity parameters show a larger increase of standard deviation for both the bedrock (average of 22% for individual and 21% for joint inversion) and CZ sub-domains (average of 86% for individual and 83% for joint inversion). Note that these results are consistent with the values of $\hat{\mathrm{R}}$; when they are small (e.g., seismic properties of the bedrock), then the variance within a given chain is similar to the variance of all chain.

In Fig. 4 we display the last posterior model realization for each of the nine MCMC chains, while in Fig. 5 we plot the corresponding vertical resistivity and velocity profiles at 50 m, 200 m and 350 m along the profile length, together with target model values. For each inversion case, the three individual MCMC chains show similar posterior model realizations and vertical profiles. Here, the consistency in the inferred interface locations is manifested by similar depths and magnitudes of the jumps in resistivity and/or velocity. Fig. 6 shows the mean, normalized standard deviation, probability maps of the in-

ferred interface location and model discrepancies (eq. 4.6). These results indicate that the probabilistic inversion of a single dataset is able to sample interfaces that are overall in good agreement with the target one. In the case of individual inversion of the DC resistivity dataset (Fig. 6a), the chains are found to sample the underlying interface well in the central area of the investigated domain (i.e., between 100 m and 300 m) but not on the sides (i.e., where there is less coverage). The probabilistic inversion of the seismic refraction dataset (Fig. 6i) samples the proper bedrock topography, but it places the interface slightly deeper than the true one. Finally the joint inversion model results (Figs. 6d and g) combine the resolution ability of the two data types and obtain an interface that is more consistent in terms of shape and depth. This is also seen by studying the discrepancies between the target models and the mean of the posterior realizations. In case of individual inversion of ERT data, Fig. 6(n) shows a thick area of over-estimated resistivity (hence the negative relative error) on the left and right side of the domain, reflecting the too shallow interface that was sampled by the MCMC chains. The mean absolute value of model discrepancy is in this case 0.24. For the individual refraction seismic inversion, Fig. 6(p) shows a layer of underestimated velocity (hence the positive relative error) along the whole profile, mimicking the target interface shape and expressing a deeper interface than the target one. This discrepancy leads to a mean absolute model discrepancy of 0.074. Finally, for both resistivity and velocity, the probabilistic joint inversion mitigates these issues. The improvement is particularly evident in the resistivity model discrepancy (Fig. 6o), where the area of over-estimation of the inferred field is significantly reduced. An improvement is also seen at 80 m and 300 m along the profile length (i.e., the valleys) for which the velocity model discrepancy (Fig. 6q) reduces significantly. The resulting mean absolute model discrepancies are 0.17 for resistivity and 0.066 for velocity. Finally, the probability maps show that the joint inversion results also increase the uncertainty on the interface location (shadow area in Fig. 6f), if compared with the individual ERT (Fig. 6c) or refraction seismic (Fig. 6m) inversion results.

Fig. 7 shows vertical profiles of the velocity and resistivity fields obtained for the different probabilistic and deterministic inversion approaches plotted together with the true model parameter values. Both the individual and joint probabilistic inversions present jumps in the physical property values at locations that are mostly in agreement with the target interface location. In case of deterministic inversion results, the physical property

values smoothly increase with depth. This results in significant underestimation of the bedrock properties.

To quantify the ability of the different inversion approaches to infer the target interface (I_{target}), we compute the average distance between I_{target} and the sampled (in case of probabilistic approaches) or inferred one (in case of deterministic inversion results):

$$\bar{D} = \frac{\int_{x_{min}}^{x_{max}} |I_{target} - I_{sampled/inferred}|}{x_{max} - x_{min}}, \tag{7}$$

where x_{min} and x_{max} are the horizontal limits of the model domain considered for the analysis of the inversion results, accordingly to the cumulative sensitivities of the methods. For the deterministic inversion, this distance is $\bar{D}$ = 9.0 m when considering the DC resistivity data and $\bar{D}$ = 6.0 m when considering the seismic refraction data. The empirical-Bayes-within-Gibbs chains infer interfaces with $\bar{D} = 7.8 \pm 0.7$ m when considering the DC resistivity and $\bar{D} = 6.2 \pm 0.6$ m when considering the seismic refraction data individually. Our proposed probabilistic joint inversion with interface constraint yields a distance of 4.1 ± 0.3 m, which is a significant improvement compared to the other inversion results.

3.2 Calhoun CZO

We now turn our attention to the results obtained when considering the field data from the Calhoun CZO. In Fig. 8, we show the (a) DC resistivity and (b) seismic refraction data together with the resulting (c-d) l_1-norm mimicking smoothness-constrained deterministic inversion results and the corresponding interface obtained using the maximum vertical gradient method. The deterministic inversion model results are plotted with transparency that is inversely proportional to the cumulative sensitivity, which shows the inability of the seismic refraction dataset to resolve subsurface properties after 400 m along the profile. For this reason, the model results shown hereafter are cut at 400 m.

For each individual or joint inversion approach, we run three independent MCMC chains for 5×10^5 iterations for which the burn-in has been reached by all inversion results. The initial models are again randomly sampled from log-uniform uncorrelated property fields. The convergence statistics of the different probabilistic inversion schemes are summarized in Table 2. Using the last 10^5 MCMC iterations, we find that the potential scale reduction factor is below 1.2 for most model parameters describing the velocity field of the bedrock sub-domain (99% of the parameters for both individual and joint inversions). However, the CZ velocity fields have only 39% (individual inversion) and 40% (joint in-

version) of the parameters with $\hat{R} < 1.2$. For the resistivity fields, only 11-12% of the CZ and 36-37% of the bedrock parameters have $\hat{R} < 1.2$ regardless if a joint or individual inversion is carried out. This implies that only an incomplete sampling of the posterior distributions were achieved for most of the model parameters considered. This suggests that model uncertainty is underestimated when using the results of one MCMC chain only. However, as for the synthetic case, this problem is partly mitigated by merging the results obtained by the three MCMC chains. For the velocity fields, the parameters included in the bedrock sub-domain have a slight augmentation in standard deviation (average of 1 % for individual and 3 % for joint inversion), while for the CZ sub-domain such increase is more significant (average of 37 % for individual and 42 % for joint inversion). In general the resistivity parameters have a larger increase of standard deviation for both bedrock (average of 46 % for individual and 47 % for joint inversion) and CZ sub-domains (average of 114 % for individual and 106 % for joint inversion).

Figure 9 presents the last posterior realization obtained for each chain and inversion case, while in Fig. 10 we show the corresponding vertical resistivity and velocity profiles at 50 m, 200 m and 350 m along the profile. The posterior model realizations obtained by inverting the DC resistivity dataset alone were already presented in *de Pasquale et al.* [2019], who stressed the similarity between the model results in terms of both posterior realizations and vertical profiles. This is similarly seen in the posterior velocity realizations obtained by inversion of the refraction seismic data alone (Figs. 9g,i,m) and in the vertical profiles (Figs. 10l-n), where the inferred interface location is manifested by velocity jumps that appear at very similar positions. Particularly, all the three chains suggest an interface that mirrors the topography, with bedrock rising almost to the surface at locations where the surface topography creates valleys at around 80 m and 350 m. The posterior models obtained by joint inversion are represented in terms of both resistivity (Figs. 9b,d,f) and velocity (Figs. 9h,i,n) fields. As for the single-dataset inversion results, the upper part of the domain is well defined for both property fields (i.e., upper 40 m). Moreover, comparing the joint inversion results with the individual inversion results (Figs. 9b,d,f with Figs. 9a,c,e and Figs. 9h,l,n with Figs. 9g,i,m), we see that the bedrock topography is mainly constrained by the seismic refraction observations. The DC resistivity data has the strongest sensitivity to heterogeneity within the bedrock layer. This is seen in the well-defined low resistivity zone of the bedrock around 300 m while the velocity fields of the bedrock are more homogeneous. Finally, around 80 m along the profile, all the poste-

rior velocity models obtained by joint inversion show a slightly deeper bedrock interface than the ones inferred from the inversion of the seismic observations alone.

In Fig. 11 we show the inversion results in terms of mean (Figs. 11a, d, g and i), normalized standard deviation (Figs. 11b, e, h and l) and probability maps of the interface locations (Figs. 11 c, f and m) for the three MCMC chains used for the individual datasets and joint inversion. The probability maps of the joint inversion results shows an increase in the uncertainty on the interface location concerning the edge of the domain, whereas at the center of the domain such uncertainty is reduced (shadow area in Fig. 11f), if compared with the individual ERT (Fig. 11c) or refraction seismic (Fig. 11m) inversion results. Finally, in Fig. 12, we plot the vertical resistivity and velocity profiles obtained from l_1-norm mimicking smoothness-constrained deterministic inversion, as well as the individual and joint inversions with interface constraints. As seen for the synthetic test case, for the probabilistic inversion results (individual ERT, seismic or joint inversion) the location of the interface is characterized by jumps in the physical property values, while in case of deterministic inversion results, the physical property values smoothly increase with depth.

4 Discussion

As an extension of the empirical-Bayes-within-Gibbs algorithm by *de Pasquale et al.* [2019], we have introduced a probabilistic joint inversion method to infer an interface separating two heterogeneous sub-domains. Our synthetic example demonstrates that the joint inversion provides more accurate interface estimates than those based on DC resistivity or seismic refraction data alone. Furthermore, a general improvement is obtained with respect to deterministic inversion results in terms of the inferred interface geometry and physical property fields, particularly with respect to bedrock properties. This is seen by studying the model discrepancies with respect to the true model for the deterministic (Fig. 3g) and probabilistic (Fig. 6n) inversions of the DC resistivity data from which it is evident that the deterministic inversion severely underestimates bedrock resistivity. For the seismic data, the deterministic inversion (Fig. 3h) over-estimates the velocity in the central area of the bedrock sub-domain (between 100 and 300 meters along the profile) while the errors for the probabilistic case (Fig. 6p) are the largest in the vicinity of the inferred interface. The improved ability to locate the interface was quantified by a measure of the distance between the target and the inferred interface (eq. 7). For the joint inversion re-

sults, this distance is at least 30 % smaller than those obtained by individual deterministic or probabilistic inversions, thereby, highlighting the added value of jointly inverting the two types of geophysical data. Despite this progress, we acknowledge that the main inherent limitation discussed at length by *de Pasquale et al.* [2019] remains. That is, in poorly resolved areas there is a significant risk that the inferred range of interface locations do not always include the true interface. This is a consequence of the stabilizing penalties that are given to the interface and the physical property variations. In short, the MCMC inversion will inherently favor smoothly varying fields, which will not necessarily include true and more variable property fields if the data are not sufficiently constraining. In the present work, we see that these problems are strongly decreased by the joint inversion. We can also not exclude that some of the sampling problems are related to the inherent poor mixing of basic Metropolis algorithms. As surveyed by *Robert et al.* [2018], there is a strong need to test and develop alternative methods that better explore the posterior landscape for a limited computational budget. For completeness, we highlight that our algorithm converges to the true interface when considering homogeneous physical properties even for the case of individual inversion of ERT data *de Pasquale et al.* [2019].

As mentioned in section 2.1, when inverting the synthetic datasets it was necessary to inflate the assumed standard deviations of the data errors from 3.7% to 10% for the ERT data and from 3 milliseconds to 5.1 milliseconds for the seismic data. This was needed to account for discrepancies in the forward model results when using different unstructured meshes in the data generation procedure (Figs. 3a and b) and in the joint inversion (Figs. 3c and d). This error inflation was calculated by computing the root mean squared errors between the datasets and the model predictions obtained from the models depicted in Figs. 3(c) and (d). Indeed, different meshes have different possible interface geometries that the MCMC chains can sample, as they are bound to follow the pre-defined node connections. The true model has a rather smooth interface and the mesh conforms to this surface, while the target interface defined by the mesh used in the MCMC forward simulations is more irregular. The travel time path calculations used are restricted to raypaths that follow the elements connecting the nodes in the network. This introduces errors in the ray geometry and it will overestimate the traveltimes [*Moser*, 1991]. We verified that such an error is not significantly reduced with the use of a refined mesh for the first arrival travel times computations. The forward algorithm used is in fact said to be *inconsistent* with the problem it seeks to approximate since it uses a network solution for

a continuous problem [*Sethian*, 1999]. An alternative that could improve the accuracy of the forward computation (and reduce the modeling error) is, for instance, to use the fast marching approach which directly approximates the solution of the underlying partial differential equation [*Sethian*, 1996; *Fomel*, 2000]. Fig. 13 shows posterior model realizations whose model predictions fit the synthetic datasets to the added data errors (i.e., without inflation to account for model errors). We see that the effects of underestimating the noise levels (ignoring the modeling error related to the different unstructured meshes) on the inversion model results are negligible for the individual probabilistic inversions of DC resistivity (Fig. 13a) and seismic (Fig. 13b) data. However, large interface artifacts appear when jointly inverting the data without this error inflation (Figs. 13c and d). The individual inversions are robust as small changes in the interface and physical properties can manifestly accommodate this model error. However, the necessary changes in the interface that accommodate this unaccounted model error are not the same for the seismic and ERT data. Indeed, a much more complex interface (i.e., with higher structure) is needed to explain both data sets when ignoring the model errors. We expect that a more detailed quantification and inclusion of model errors [*Hansen et al.*, 2014] would further improve our results. We leave this topic for future research. This problem does not appear as acute for the Calhoun CZO inversions, probably because of the rather conservative choice of the data error models (i.e., the noise level considered for both the datasets is rather large).

The results obtained by jointly inverting the Calhoun datasets (Figs. 11d,f,g and 12) indicate sharp interfaces at locations that agree well with the more gradual transitions presented by *St. Clair et al.* [2015]. In agreement with *St. Clair et al.* [2015], the transition between the CZ and bedrock has a shape that mirrors surface topography, with bedrock rising almost to the surface below the valleys. Moreover, all the inversion results clearly indicate two regions of lower resistivities within the bedrock domain that are located below the topographic depressions (i.e., around 80 m and 350 m along the profile; Figs. 8c, 9a-f, 10c and f, 11a and d, 12c). Extensive testing (not shown) indicates that it is impossible to fit the ERT data without introducing such lower-resistivity regions. We also note that corresponding features are also visible in Fig. 4a in *St. Clair et al.* [2015]. These lower-resistivity zones are likely to be caused by a higher fracture intensity. In Fig. 14, we display the resistivity and velocity values corresponding to the bedrock at the inferred interface between bedrock and regolith. The resistivity values drop below the topographic depressions, while the seismic velocities slightly increase. We attribute the higher seismic

velocity to model errors (as explained above) that is likely to lead to over-estimated velocities that compensate for ray-paths that are too long. Since there is no seismic evidence of mechanical weakening in the bedrock below the valleys, we expect that the decrease in resistivity is caused by a small increase in fracture porosity and that these fractures may be clay-filled. Induced polarization data or a borehole would shed more light about the presence of clay-filled fractures and lead to more conclusive findings.

5 Conclusions

The geometry and depth to the regolith-bedrock interface impact many surface and near-surface processes. *de Pasquale et al.* [2019] presented a probabilistic inversion framework to infer the posterior distribution of an interface separating two heterogeneous subdomains. We have extended this framework to joint inversion of DC resistivity and seismic refraction datasets in presence of a common interface with unknown geometry. The synthetic example demonstrates that the interface inferred by joint inversion is more accurate than those obtained by individual deterministic or probabilistic inversions of either DC resistivity or seismic refraction data. When applied to field-data from the Calhoun CZO in South Carolina, we find in agreement with *St. Clair et al.* [2015] that the bedrock mirrors the surface topography. Furthermore, the sampled posterior resistivity fields suggests that bedrock below the valleys in surface topography has an enhanced fracture porosity. Since there is no evidence of a coincident mechanically-weak zone in seismic P-wave velocity, we suggest that the fractures are clay-filled. Our methodology can be implemented to case studies where other geophysical measurements are used and extended to the inversion of three or more datasets.

6 Acknowledgements

This research was supported by the Swiss National Science Foundation under grant 200021-155924. We thank Joseph Doetsch and Florian Wagner for their precious inputs concerning the use of the pyBERT and pyGIMLI libraries. We are grateful to Steven Hoolbrook who kindly provided the data and shared some early comments related to this work. We thank Editor Mark Everett, Florian Wagner and two anonymous reviewers for their many useful comments.

References

Anderson, S.P., Dietrich, W.E. & Brimhall, G.H., 2002. Weathering profiles, mass-balance analysis, and rates of solute loss: Linkages between weathering and erosion in a small, steep catchment. *Geological Society of America Bulletin* 114, 1143–1158.

Archie. G.E., 1942. The electrical resistivity log as an aid in determining some reservoir characteristics. *Transactions of the Metallurgical Society of AIME* 164, 54–62.

Barton, N., 2009. *Rock Quality, Seismic Velocity, Attenuation and Anisotropy*. Environmental & Engineering Geoscience. Taylor & Francis, UK.

Berge, P.A., Berryman, J.G., Bertete-Aguirre, H., Bonner, P., Roberts, J.J., & Wildenschild, D., 2000. *Joint inversion of geophysical data for site characterization and restoration monitoring*. LLNL Rep. UCRL-ID-128343, Proj. 55411, Lawrence Livermore Natl. Lab., Livermore, Calif.

Binley, A.& Kemna, A., 2005. DC resistivity and induced polarization methods. *In: Rubin Y., Hubbard S.S. (eds). Hydrogeophysics.* Water Science and Technology Library, vol 50. Springer, Dordrecht.

Biot, M.A., 1956 Theory of propagation of elastic waves in a fluid-saturated porous solid. I. low frequency range and II. higher frequency range. *Journal of Acoustical Society of America* 28, 1809–1831.

Brantley, S.L., Goldhaber, M.B. and Ragnarsdottir, K.V., 2007. Crossing disciplines and scales to understand the critical zone. *Mineralogical Society of America* 3(5), 307–314.

Brown, E.T, 1981. *Rock characterization, testing & monitoring: ISRM suggested methods.* Pergamo Press, Oxford.

Chambers, J.E., Wilkinson, P.B., Wardrop, D., Hameed. A., Hill, I., Jeffrey, C., Loke, M.H., Meldrum, P.I., Kuras, O., Cave, M. & Gunn, D.A., 2012. Bedrock detection beneath river terrace deposits using three-dimensional electrical resistivity tomography. *Geomorphology* 177-178, 17–25.

Chapman, M., 2003. Frequency–dependent anisotropy due to meso-scale fractures in the presence of equant porosity. *Geophysical Prospecting* 51, 369–379.

Dahlin, T., 1996. 2D resistivity surveying for environmental and engineering applications. *First Break* **14**(7), 275–283.

de Pasquale, G., Linde, N., Doetsch, J. & Holbrook, W., 2019. Probabilistic inference of subsurface heterogeneity and interface geometry using geophysical data. *Geophysical*

Journal International. **217**(2), 816–831.

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

Dijkstra, E.W., 1959. A note on two problems in connection with graphs. *Numerische Mathematik* **1**, 269-271.

Doetsch, J., Linde, N., Pessognelli, M., Green, A. & Günter, T., 2012. Constraining 3-D electrical resistance tomography with GPR reflection data for improved aquifer characterization. *Journal of Applied Geophysics* 78, 68–76.

Fomel, S., 2000. A variational formulation of the fast marching eikonal solver. *Stanford Exploration Project, Report SERGEY* 455-476.

Gallardo, L.A. & Meju, M.A., 2004. Joint two-dimensional DC resistivity and seismic travel time inversion with cross-gradients constraints *Journal of Geophysical Research* 109, 1–11.

Gassmann, F., 1951. Über die Elastizität poröser medien. Vierteljahresschrift der naturforschenden. *Gesellschaft in Zurich* 96, 1–23.

Geweke, J., 1992. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, Eds., *Bayesian Statistics, Vol. 4* Clarendon Press, Oxford, pp. 169-193.

Gelman, A. & Rubin, D.B., 1992. Inference from iterative simulation using multiple sequences. *Statistical Science* **7**, 457-511.

Günther, T., Rücker, C. & Spitzer, K., 2006. Three-dimensional modelling and inversion of dc resistivity data incorporating topography-II. Inversion. *Geophysical Journal International* **166**, 506-517.

Hansen, T.M., Cordua, K.S., Jacobsen, B.O. & Mosegaard, K., 2014. Accounting for imperfect forward modeling in geophysical inverse problems – Exemplified for crosshole tomography. *Geophysics* 79(3), H1–H21.

Heimsath, A.M., Dietrich, W.E., Nishiizumi, K. & Finkel, R.C., 1997. The soil production function and landscape equilibrium. *Soil Science Society of America Journal* 57(1), 256–261.

Hellman, K., Ronczka, M., Günther, T., Wennermark, M., Rü, C. & Dahlin, T., 2017 Structurally coupled inversion of ERT and refraction seismic data combined with cluster-based model integration. *Journal of Applied Geophysics* 143, 169–181.

Hirsch, M., Bentley, L. & Dietrich, P., 2007. A comparison of electrical resistivity, ground penetrating radar and seismic refraction results at a river terrace site. *Journal of Environmental and Engineering Geophysics* 13, 325–333.

Hudson, J.A., 1981. Wave speeds and attenuation of elastic waves in material containing cracks. *Geophysical Journal of Royal Astronomical Society* 64, 133–150.

Hsu, H.-L., Yanites, B.J., Chen, C.-c & Chen, Y.-G., 2010. Bedrock detection using 2D electrical resistivity imaging along the Peikang River, central Taiwan. *Geomorphology* 114, 406–414.

Jones, D.P. & Graham, R.C., 1993. Water-holding characteristics of weathered granitic rock in chaparral and forest ecosystems. *Nature* 388, 358–361.

Juhojuntti, N. & Kamm, J., 2015. Joint inversion of seismic refraction and resistivity data using layered models âĂŤ Applications to groundwater investigation. *Geophysics* **80**(1), EN43–EN55.

Kozlovskaya, E., Vecsey, L., Plomerová, J. & Raita1, T., 2007. Joint inversion of multiple data types with the use of multiobjective optimization: problem formulation and application to the seismic anisotropy investigations. *Geophysical Journal International* 171, 761–779.

Lanni, C., McDonnell, J., Hopp, L. & Rigon, R., 2013. Simulated effect of soil depth and bedrock topography on near surface hydrologic response and slope stability. *Earth Surface Processes and Landforms* 38, 146–159.

Lee, S.G. & Freitas, M.H., 1990. Seismic refraction surveys for predicting the intensity and depth of weathering and fracturing in granitic masses. In:*Field Testing in Engineering Geology*. Geological Society of London.

Lesmes, D.P. & Friedman, S.P., 2005. Relationships between the electrical and hydrogeological properties of rocks and soils. *In: Rubin Y., Hubbard S.S. (eds). Hydrogeophysics.* Water Science and Technology Library, vol 50. Springer, Dordrecht.

Linde, N., Binley, A., Tryggvason, A., Pedersen, L.B. & Revil, A., 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. & Doetsch, J., 2016. Joint inversion in hydrogeophysics and near–surface geophysics. *In: Moorkamp M., Lelievre P., Linde N., and Khan A. (eds). Integrated Imaging of the Earth: Theory and Applications.* American Geophysical Union.

Linder, S., Paasche, H., Tronicke, J., Niederleithinger, E. & Vienken, T., 2010. Zonal cooperative inversion of crosshole P-wave, S-wave and georadar traveltime data sets. *Journal of Applied Geophysics* 72(4), 254–262.

Liu, E. and Martinez, A., 2012. *Seismic Fracture Characterization. Concepts and Practical Applications.* Houten, The Netherlands: EAGE Publications bv.

Lowry, T., Allen, M.B. & Schive, P.N., 1989. Singularity removal: a refinement of resistivity modelling techniques. *Geophysics* **41**, 766-774.

Mavko, G., Dvorkin, J., & Mukerji, T., (2009). *Rock Physics Handbook: Tools For Seismic Analysis in Porous Media.* Cambridge: Cambridge University Press.

Meyn, S.P. & Tweedie, R.L., 1993. *Markov Chains and Stochastic Stability.* Springer-Verlag, London UK.

Moser, T.J., 1991. Shortest path calculation of seismic rays. *Geophysics* **56**(1), 59-67.

Nur, A & Simons, G., 1969. The effect of saturation on velocity in low porosity rocks. *Earth Planet Science Letters* 7, 183–193.

Olona, J., Pulgar, J.A., Fernández-Vejo, G., López-Fernández, C. & González-Cortina, J.M., 2010. Weathering variations in a granitic massif and related geotechnical properties through seismic and electrical resistivity methods. *Near Surface Geophysics* 8, 585–599.

Onda, Y., Tsujimura, M.& Tabuchi, H., 2004. The role of subsurface water flow paths on hillslope hydrogeological processes, landslides and landform development in steep mountains of Japan. *Hydrological Processes* 18(4), 637–650.

Rempe, D.M. & Dietrich, W.E., 2014. A bottom-up control on fresh-bedrock topography under landscapes. *Proceedings of the National Academy Society* 111(18), 6576–6581.

Robert, C. P., Elvira, V., Tawn, C. & Wu, C., 2018. Accelerating MCMC algorithms. *Wiley Interdisciplinary Reviews: Computational Statistics* 10, e1435.

Reneau, S.L. & Dietrich, W.E., 1987. Size and location of colluvial landslides in a steep forested landscape. *Earth Surface Processes and Landforms* 38, 146–159.

Rosas-Carbajal, M., Linde, N., Kalscheuer, T. & Vrugt, J.A., 2014. Two-dimensional probabilistic inversion of plane-wave electromagnetic data: methodology, model constraints and joint inversion with electrical resistivity data. *Geophysical Journal International*, 196, 1508–1524.

Rücker, C., Günther, T. & Spitzer, K., 2006. Three-dimensional modelling and inversion of dc resistivity data incorporating topography-I. Modelling. *Geophysical Journal Interna-*

tional **166**, 495-505.

Rücker, C., Günther, T. & Wagner, F.M., 2017. pyGIMLi: An open-source library for modelling and inversion in geophysics. *Computers and Geosciences* 109, 106–123.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics* 40, 3.1–3.29.

Sass, O., 2007. Bedrock detection of talus thickness assessment in the European Alps using geophysical methods. *Journal of Applied Geophysics* 62, 254–269.

Schmitt, D., 2015. Geophysical properties of the near surface Earth: Seismic properties. *In: Schubert, G. Treatise on Geophysics. 11 Volume II.* Elsevier, Amsterdam NL.

Schrott, L., Hördt, A. & Dikau, R., 2003. *Geophysical applications in geomorphology.* Zeitschrift für Geomorphologie. Supplementband 132, V–VII.

Sethian, J.A., 1996. A fast marching level set method for monotonically advancing fronts. *Proceedings of the National Academy of Sciences*, 93(4), 1591–1595.

Sethian, J.A., 1999. Fast marching methods *SIAM Review* 41(2), 199–235.

Si, H., 2003. TetGen, a Delaunay-Based Quality Tetrahedral Mesh Generator. *ACM Transactions on Mathematical Software (TOMS)* **41**(2), Article 11, 36 pages.

St. Clair, J., Moon, S., Holbrook, W.S., Perron, J.T., Riebe, C.S., Martel, S.J., Carr, B., C. Harman, C., Singha, K. & Richter, B. deB., 2015. Geophysical imaging reveals topographic stress control of bedrock weathering. *Science* 350(6260), 534–537.

Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation* Society of Industrial and Applied Mathematics, Philadelphia USA.

Tye, A., Kessler, H., Ambrose, K., Williams, J.D.O., Tragheim, D., Scheib, A., Raines, M. & Kuras, O., 2011. Using integrated near-surface geophysical surveys to aid mapping and interpretation of geology in an alluvial landscape within a 3D soil-geology framework. *Near Surface Geophysics* 9, 15–31.

Wyllie, M.R.J., Gregory, A.R. & Gardener, G.H.F, 1956. Elastic wave velocities in heterogeneous and porous media. *Geophysics* 21, 41–70.

Wyllie, M.R.J., Gregory, A.R. & Gardener, G.H.F, 1958. An experimental investigation of factors affecting elastic wave velocities in porous media. *Geophysics* 23, 459–493.

Xu, S., 1998. Modelling the effects of fluid communication on velocities in anisotropic porous rocks. *International Journal of Solids & Structures* 35, 4685–4707.

Zhou, H.-W., 2014. *Practical Seismic Data Analysis.* Cambridge University Press.

7 Tables and Figures

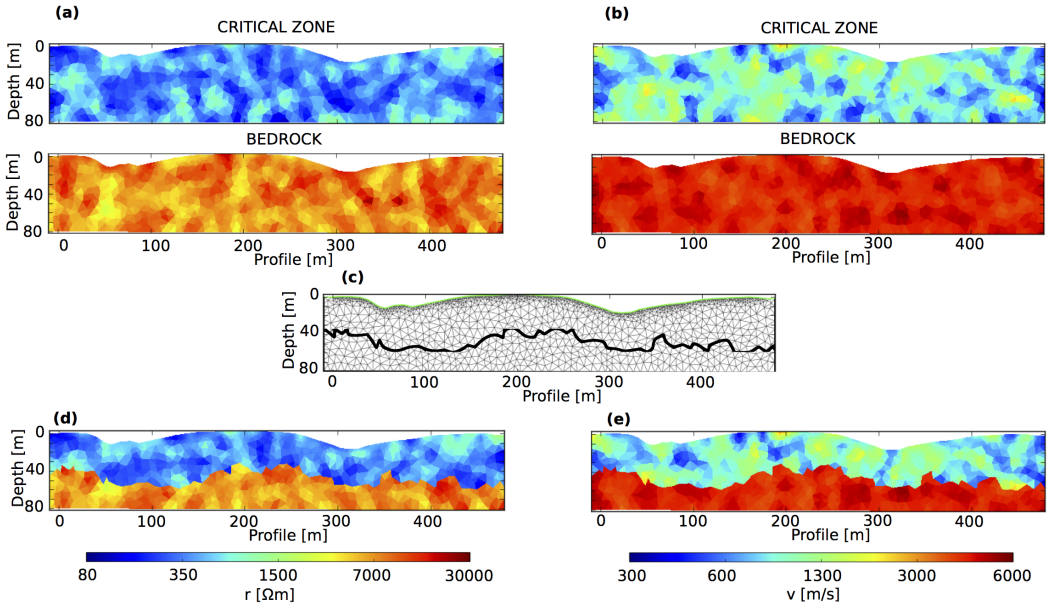

Figure 1. Proposed model setting for joint inversion of DC resistivity and refraction seismic datasets in the presence of an unknown interface (**I**), which divides the investigated subsurface into two sub-domains. (a) Electrical resistivity fields of the CZ and bedrock. (b) Seismic P-wave velocity fields of the CZ and bedrock. (c) The interface, defined by connected nodes in the forward computation mesh, is used to assign (d) the electrical resistivity ($\mathbf{m}^r$) and (e) P-wave velocity ($\mathbf{m}^v$) used for forward calculations.

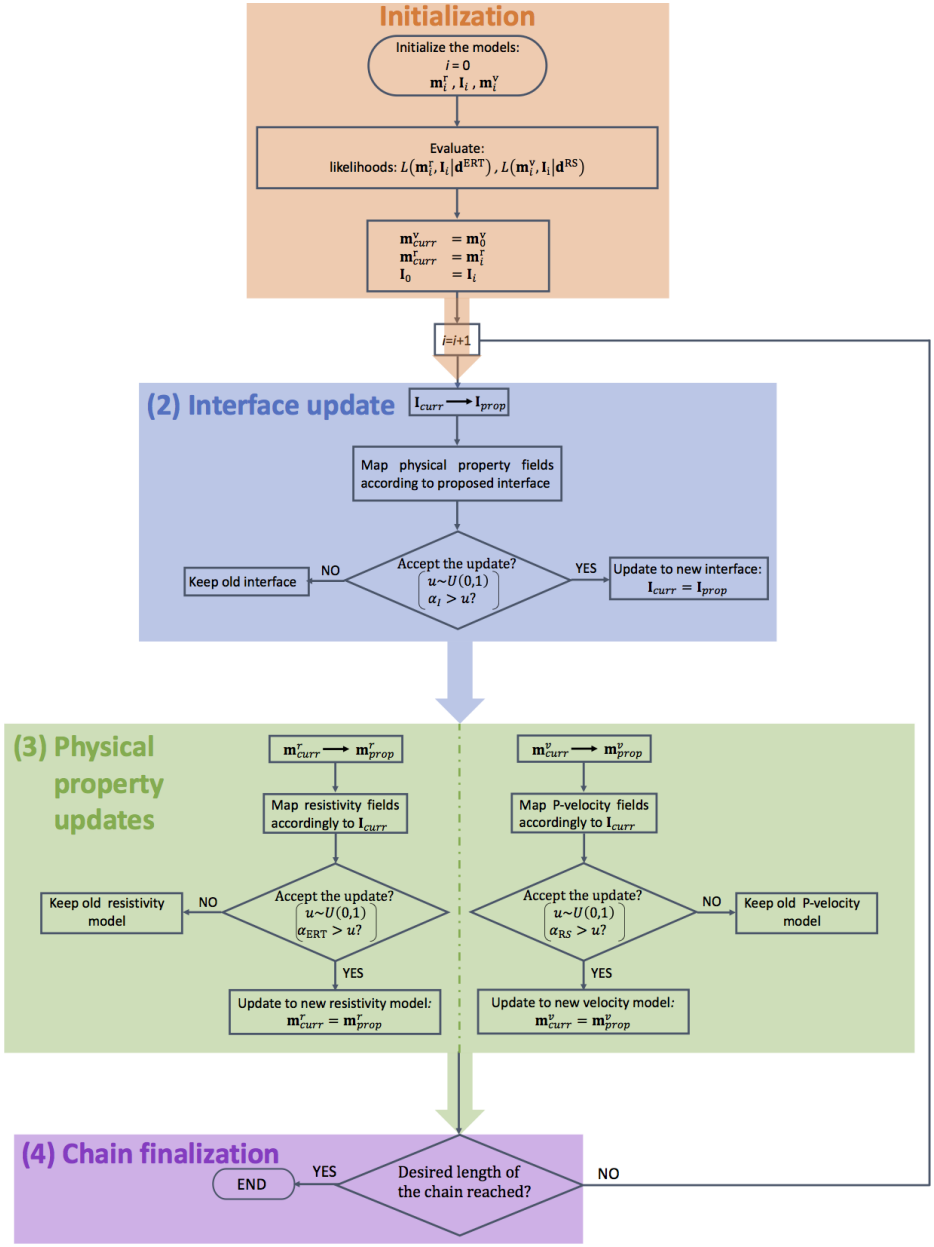

Figure 2. Schematic flow chart of the proposed probabilistic joint inversion algorithm with a common interface. First, we initialize the electrical resistivity ($\mathbf{m}_0^r$) and seismic P-wave velocity ($\mathbf{m}_0^v$) models (as described in Fig. 1), evaluate their likelihoods (L) and the corresponding structural constraint functions. We then sequentially evaluate proposed interface and physical property updates in an iterative scheme until the chosen length of the chain is reached. Here, $u \sim U(0,1)$ represents a randomly drawn number with uniform probability between 0 and 1, while the acceptance probabilities α_I, α_{ERT} and α_{RS} are described in eqs. 1, 2 and 3, respectively. For more details, please refer to *de Pasquale et al.* [2019].

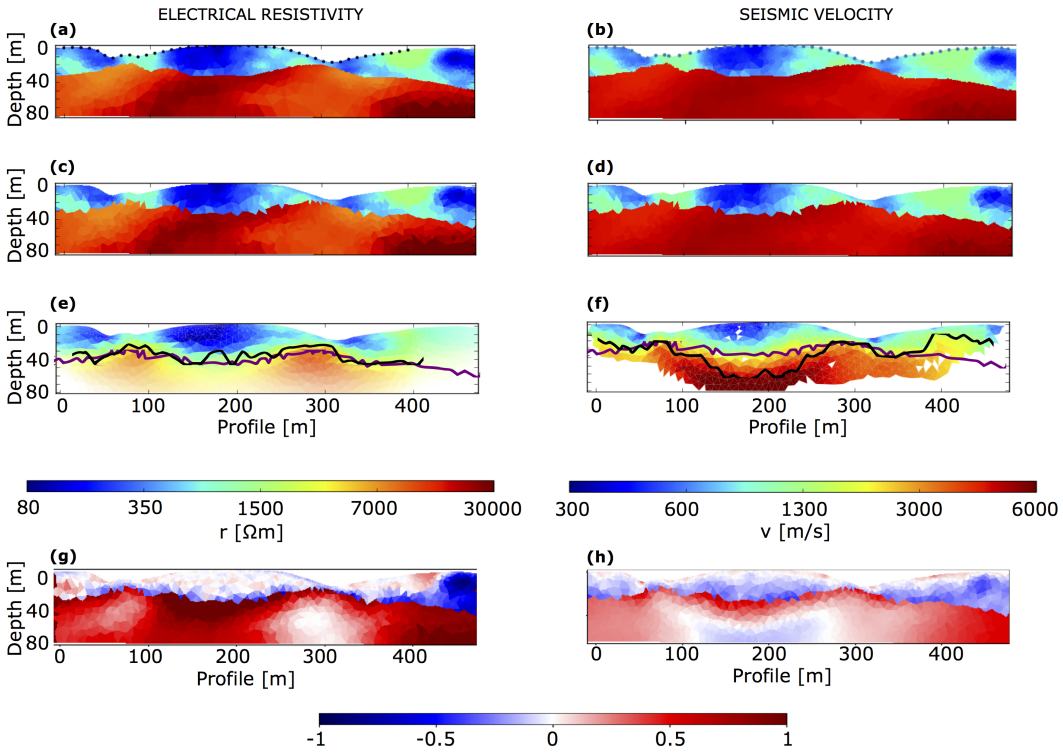

Figure 3. (a) Resistivity model and electrode positions (black dots) used to generate the synthetic DC resistivity dataset and (c) interpolation of (a) on the unstructured mesh used for ERT forward modeling. (b) Velocity model and shot positions (blue dots) used to generate the synthetic refraction seismic dataset and (d) interpolation of (b) on the unstructured mesh used for the refraction seismic forward modeling. Deterministic inversion results based on the (e) DC resistivity and (f) seismic refraction observations. In (e) and (f) the transparency is inversely proportional to the cumulative sensitivity, the black line represents the interface obtained by the maximum vertical gradient method and the purple line is the interface we aim to infer. Model discrepancies for deterministic inversion of the (g) DC resistivitiy and (h) seismic refraction observations.

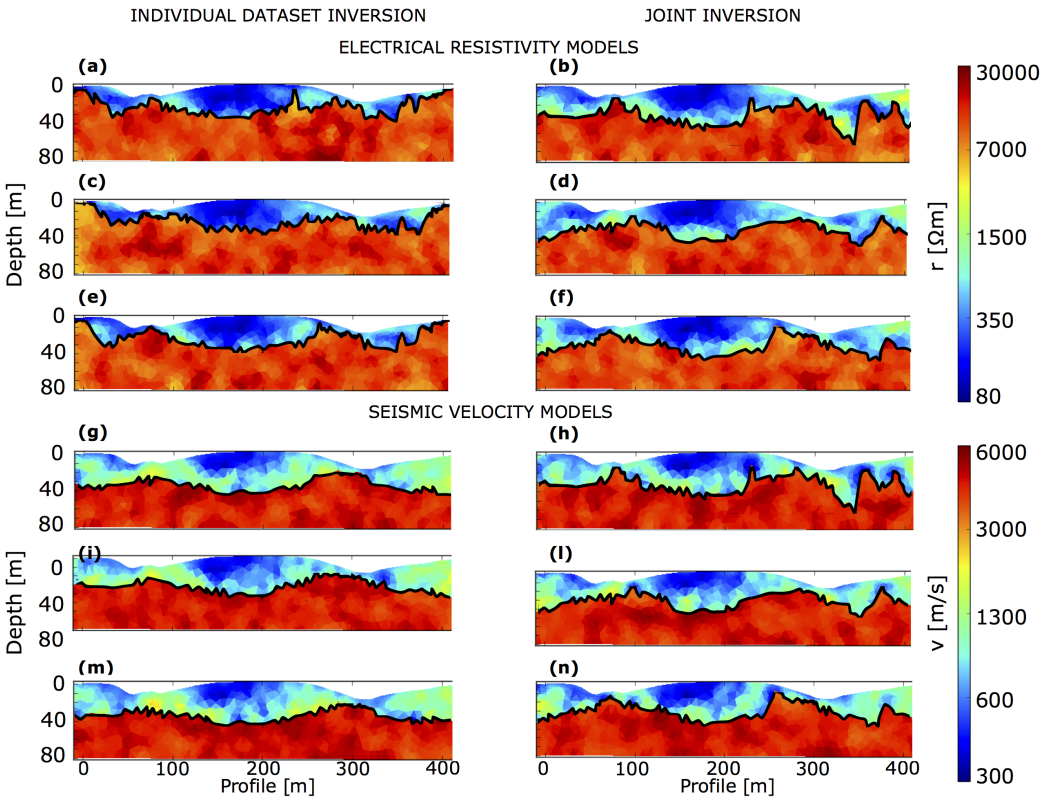

Figure 4. Last posterior model realizations from probabilistic inversion of synthetic datasets. (a), (c) and (e) represent resistivity model realizations from each MCMC chain for individual ERT inversion, while (b), (d) and (f) show such realizations in case of joint inversion. (g), (i) and (m) are P-velocity model realizations from each MCMC chain for individual refraction seismic inversion, while (h), (l) and (n) show such realizations in case of joint inversion. The inferred interface of each realization is represented with a black line.

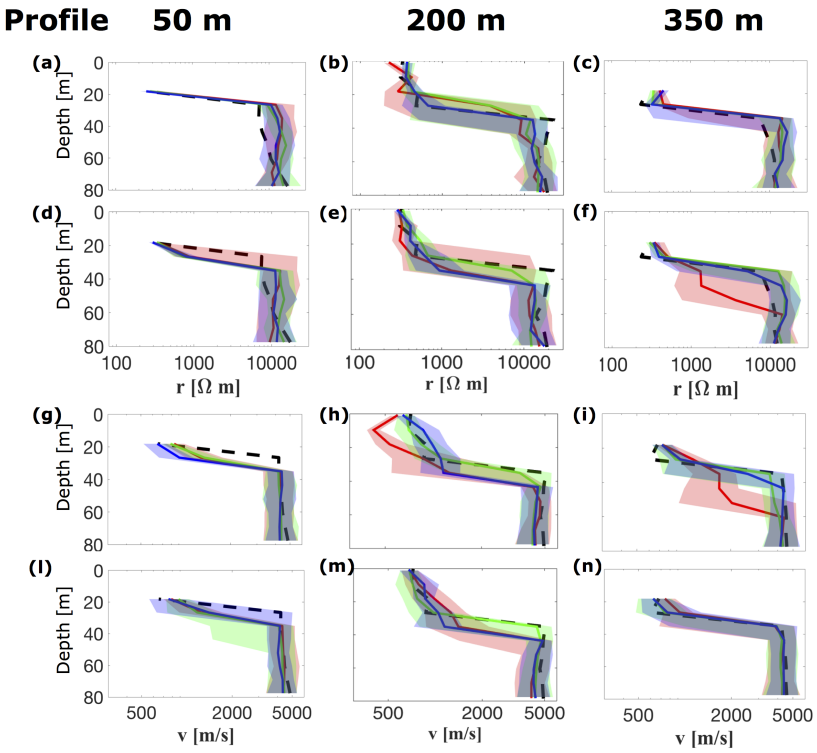

Figure 5. Vertical resistivity and P-wave velocity profiles obtained by probabilistic inversion of the synthetic datasets. Blue, red and green indicate the mean model of each chain (solid lines) with the inferred posterior ranges (shadow areas) and the black dashed lines represent the target model. (a)-(c) Inferred resistivity values from inversion of DC resistivity data. (d)-(f) Inferred resistivity and (g)-(i) velocity values from joint inversion. (l)-(n) Inferred velocity values from inversion of individual seismic refraction data.

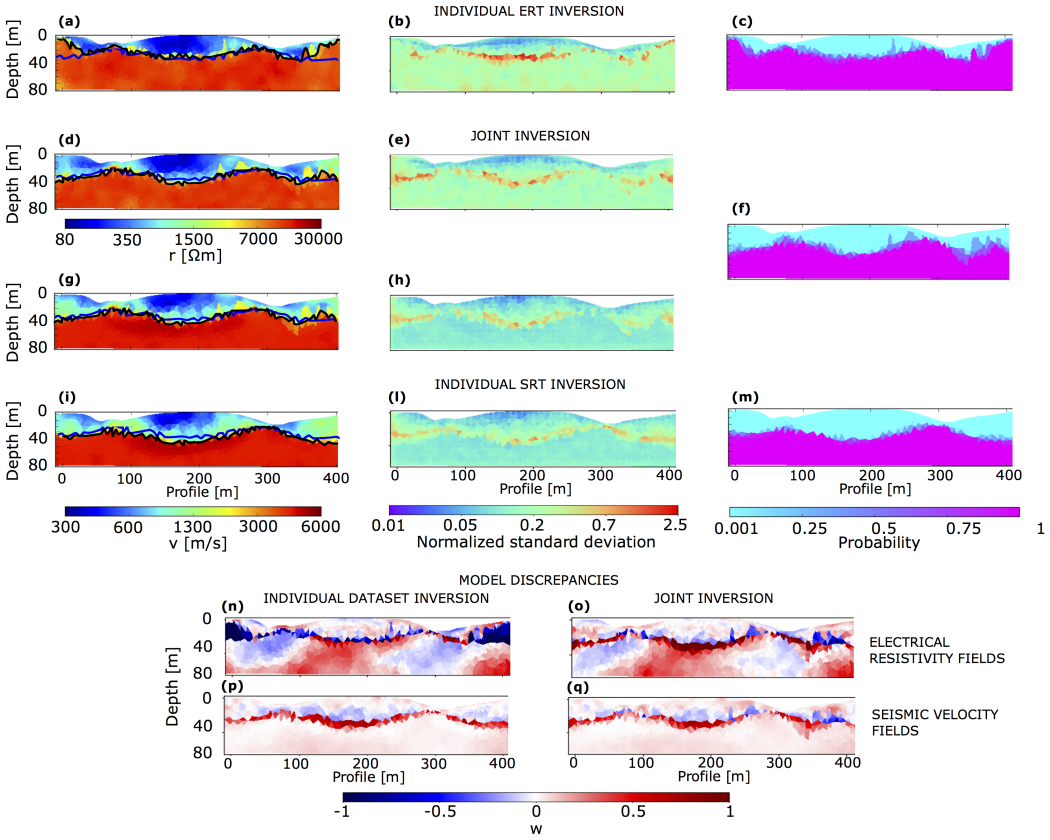

Figure 6. Summary of posterior realizations for the three MCMC chains used for each synthetic inversion case. Mean of the posterior model realizations for (a) individual DC resistivity and (i) seismic refraction inversions; (d) and (g) for joint inversion. The mean inferred interfaces are represented with a black line, while the interface we aim to infer is shown in blue. Standard deviation of the resistivities, divided by the corresponding mean values of (b) individual DC resistivity inversion and (e) joint inversion. Standard deviation of the velocities, divided by the corresponding mean values for probabilistic inversion of (h) individual refraction seismic data and (l) joint inversion. Probability for each cell of the forward mesh to be part of the bedrock in case of probabilistic inversion of individual (c) DC resistivity (m) and refraction seismic data and in case of (f) joint inversion. Model discrepancies for the resistivity fields inferred from probabilistic inversion of (n) individual ERT data and (o) joint inversion. Relative model errors for the velocity fields inferred from probabilistic inversion of (p) individual seismic refraction data and (q) joint inversion.

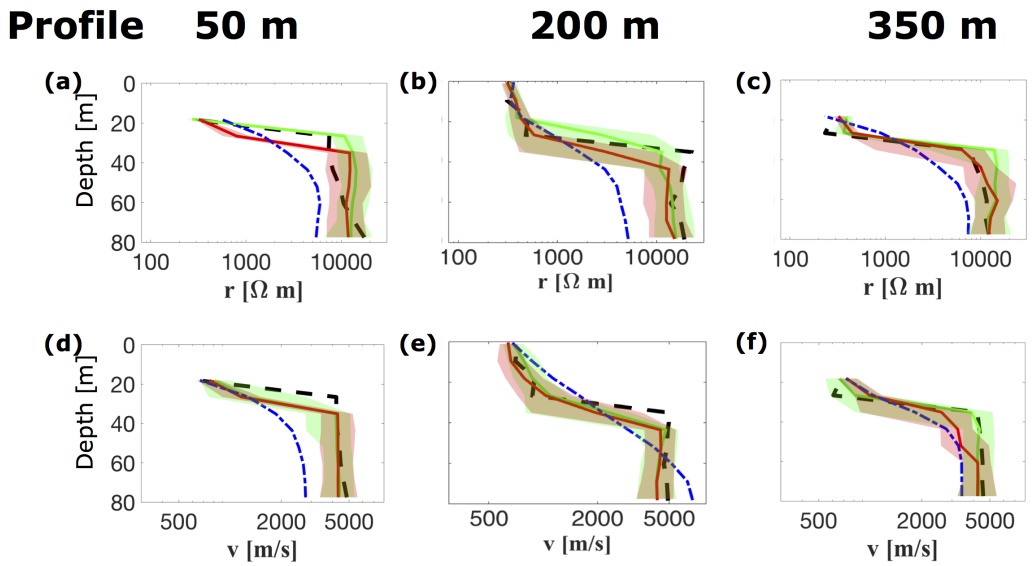

Figure 7. Vertical resistivity and P-wave velocity profiles from inversion of the synthetic datasets. Red solid lines represent the mean of the probabilistic joint inversion results and the red shadow areas the inferred posterior ranges, the green solid lines with the green shadow areas represent individual dataset probabilistic inversion results, the blue dotted lines are obtained from deterministic inversion and the black dashed lines are the target profiles. (a)-(c) Inferred resistivity values and (d)-(f) inferred velocity values.

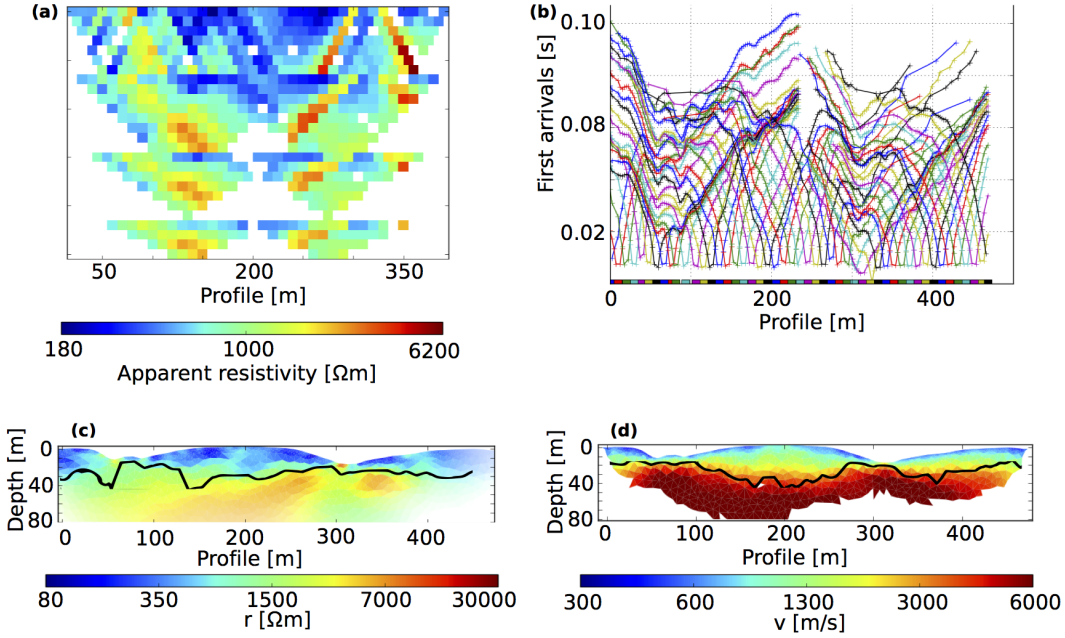

Figure 8. (a) ERT observations and (b) refraction seismic first-arrival travel times from the Calhoun CZO field site. Deterministic inversion results based on (c) DC resistivity and (d) seismic refraction observations. In (c) and (d) the models transparency is inversely proportional to the cumulative sensitivity and the black line represents the interface obtained by the maximum vertical gradient method.

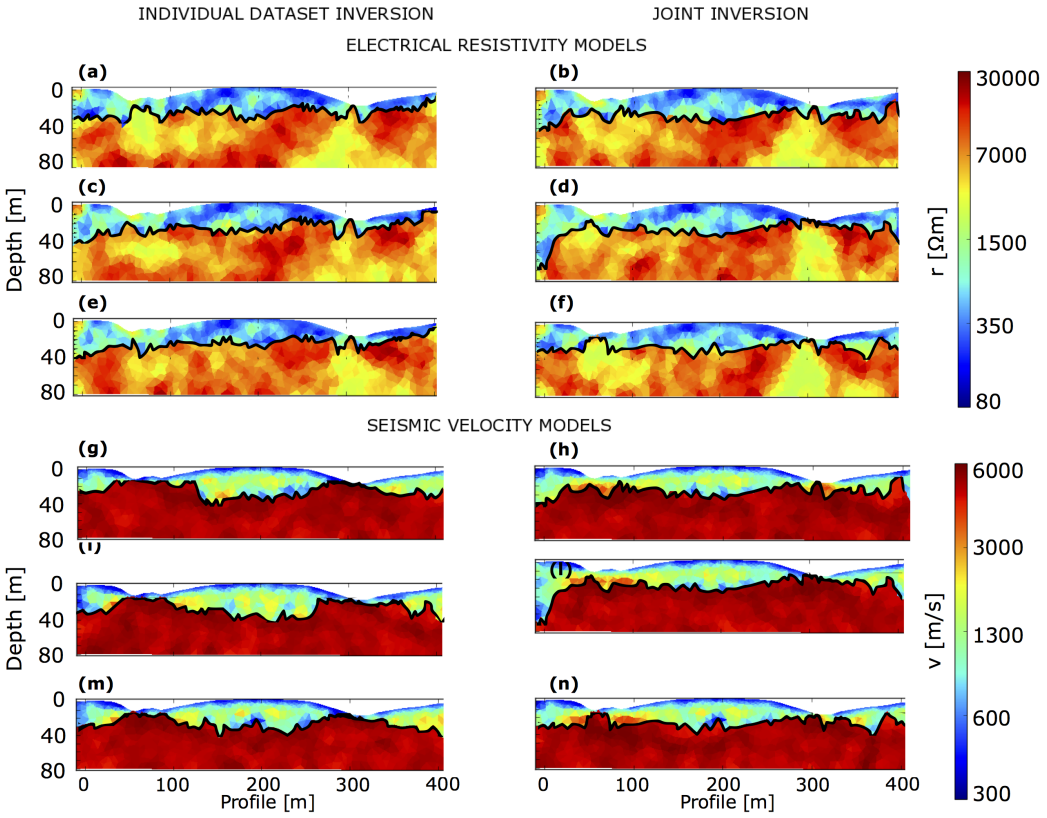

Figure 9. Last posterior model realizations from probabilistic inversion of the Calhoun datasets. (a), (c) and (e) represent resistivity model realizations from each of the MCMC chains for individual DC resistivity inversion, while (b), (d) and (f) show such realizations in case of joint inversion. (g), (i) and (m) are P-velocity model realizations from each of the MCMC chains for individual refraction seismic inversion, while (h), (l) and (n) show such realizations in case of joint inversion. The inferred interface of each realization is represented with a black line.

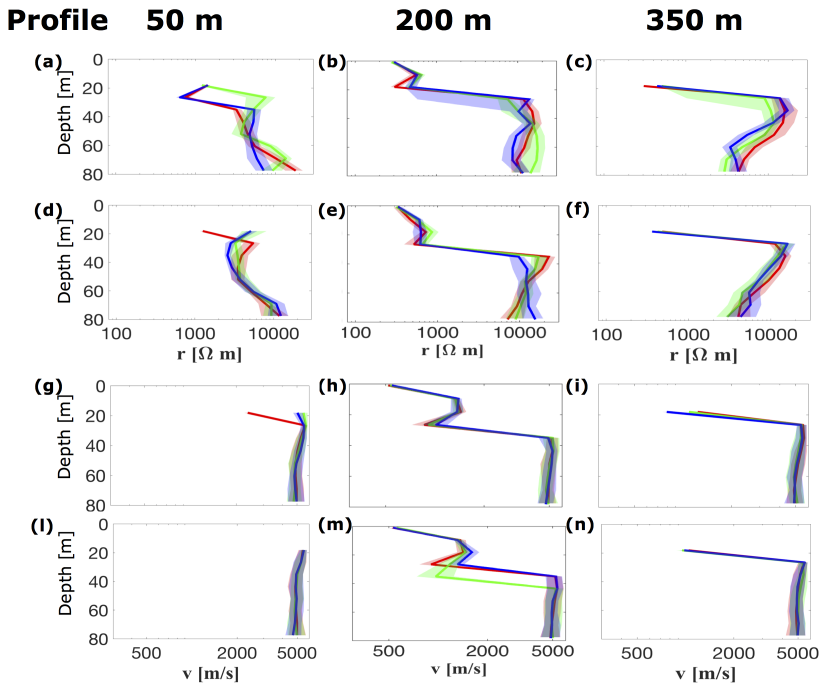

Figure 10. Vertical resistivity and P-wave velocity profiles from probabilistic inversion of the Calhoun datasets. Blue, red and green indicate the mean model of each chain (solid lines) with the inferred posterior ranges (shadow areas). (a)-(c) Inferred resistivity values from inversion of DC resistivity data alone. (d)-(f) Inferred resistivity and (g)-(i) velocity values from joint inversion. (l)-(n) Inferred velocity values from inversion of the refraction seismic dataset alone.

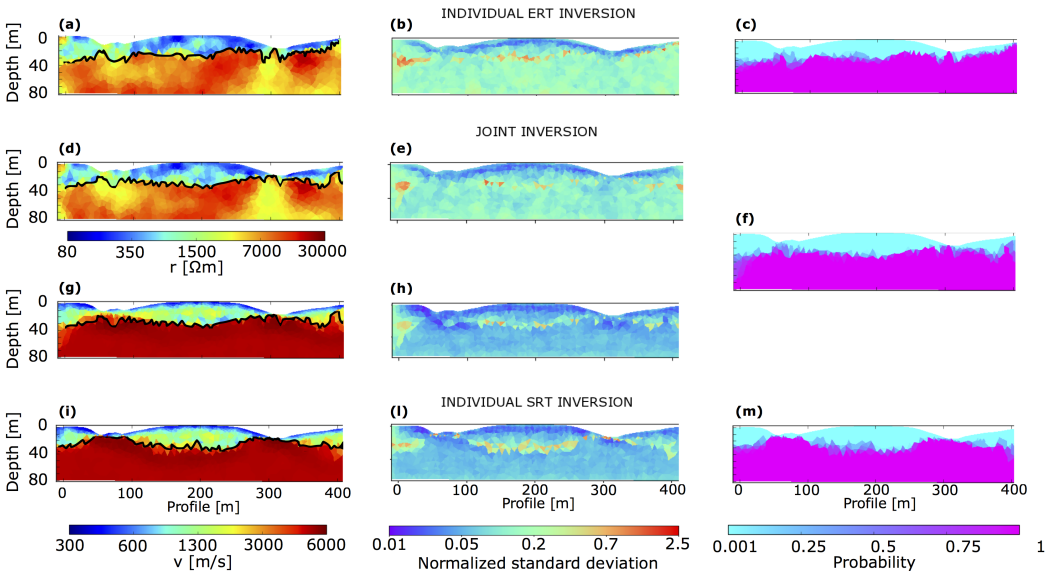

Figure 11. Summary of posterior realizations for the three MCMC chains used for the inversion of the Calhoun datasets. Mean of the posterior model realizations when inverting for individual (a) DC resistivity and (i) refraction seismic dataset; (d) and (g) when jointly inverting the two datasets. The mean inferred interfaces are represented with a black line. Standard deviation of the resistivities, divided by the corresponding mean values for probabilistic inversion of (b) individual DC resistivity data and (e) for joint inversion. Standard deviation of the velocities, divided by the corresponding mean values for probabilistic inversion of the (h) individual refraction seismic dataset and (l) for joint inversion of DC resistivity data and refraction seismic dataset. Probability for each cell of the forward mesh to be part of the bedrock in case of probabilistic inversion of individual (c) DC resistivity, (m) refraction seismic data (m) and (f) joint inversion.

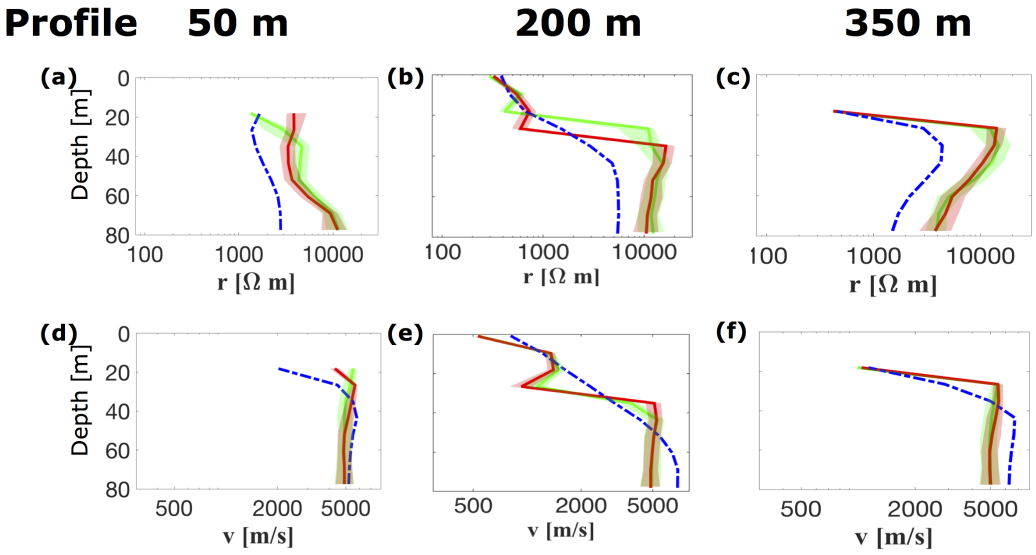

Figure 12. Vertical resistivity and P-wave velocity profiles obtained by inversion of the Calhoun datasets. Red solid lines are used to represent the mean of probabilistic joint inversion results and the red shadow areas the inferred posterior ranges, the green solid lines with the green shadow areas represent individual dataset probabilistic inversion results and the blue dotted lines are obtained from deterministic inversion. (a)-(c) Inferred resistivity values and (d)-(f) Inferred velocity values.

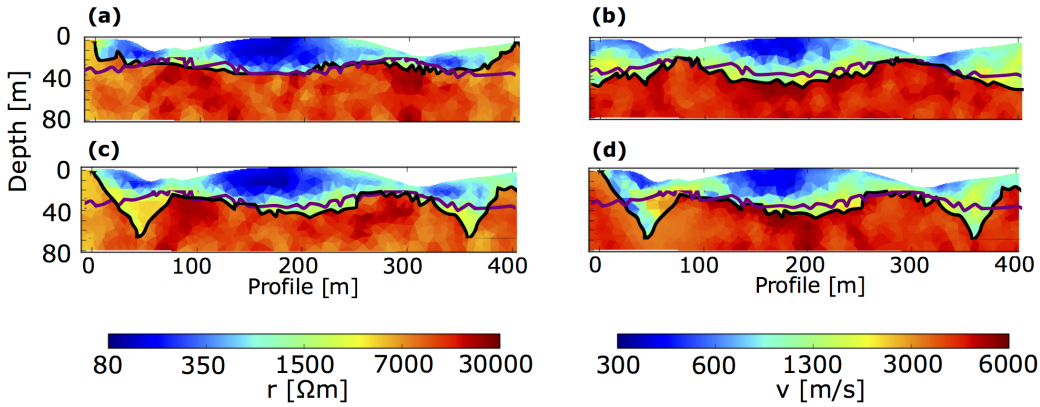

Figure 13. Posterior model realizations from probabilistic inversion of synthetic datasets when the model errors are not taken into account. (a) Resistivity model realization inferred from inversion of individual DC resistivity data and (b) velocity model realizations inferred from inversion of individual refraction seismic data. (c) Resistivity and (d) velocity model realizations inferred from joint inversion. The inferred interface of each realization is represented with a black line and the target interface with a purple line.

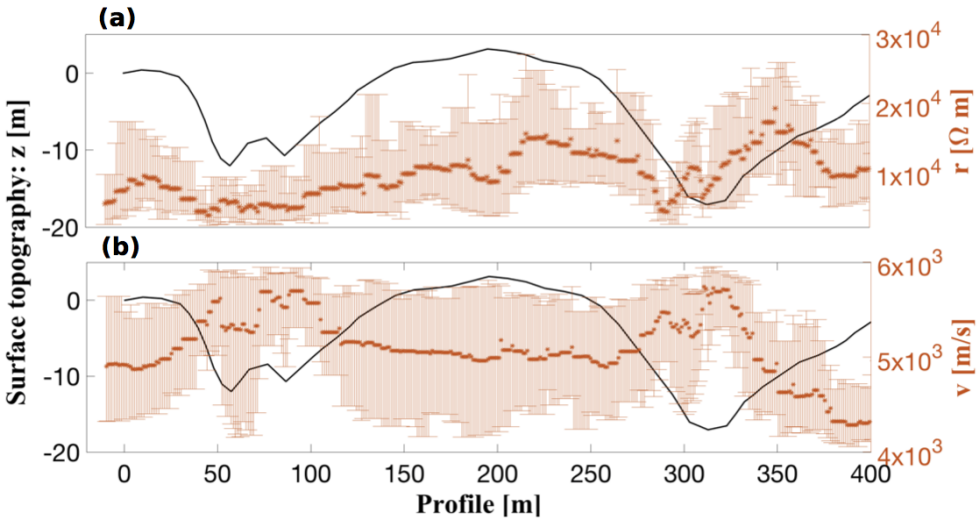

Figure 14. Inferred mean values and ranges of (a) resistivity and (b) velocity at the interface location as obtained from probabilistic joint inversion of the Calhoun datasets. In both plots, the black line represents surface topography.

Table 1. Convergence statistics for the probabilistic inversion of synthetic datasets. ERT stands for probabilistic inversion of DC resistivity data with the empirical-Bayes-within-Gibbs approach, RS refers to empirical-Bayes-within-Gibbs inversion of the seismic refraction data and Joint refers to the probabilistic joint inversion of the two datasets with interface constraints. The sub-scripts (1,2,3) indicate the three different chains used for each inversion routine. AR stands for the averaged acceptance rate of the chains after burn-in and the sub-scripts I, v and r refer respectively to the interface, velocity and resistivity fields. Finally the burn-in columns enumerate the number of iterations needed for each chain to converge to the target posterior distribution [Geweke, 1992].

Chain	AR_I [%]	AR_v [%]	AR_r [%]	Burn-in$_v$ [n°of iterations]	Burn-in$_r$ [n°of iterations]
ERT_1	7.9	-	18.3	-	2.0×10^5
ERT_2	10.3	-	18.2	-	2.2×10^5
ERT_3	7.9	-	18.3	-	2.1×10^5
RS_1	4.9	20.7	-	0.9×10^5	-
RS_2	5.1	20.7	-	1.0×10^5	-
RS_3	5.6	21.0	-	0.8×10^5	-
$Joint_1$	6.6	20.5	18.6	1.1×10^5	2.4×10^5
$Joint_2$	5.2	20.5	18.8	1.0×10^5	2.5×10^5
$Joint_3$	7.2	20.6	18.5	1.2×10^5	2.3×10^5

Table 2. Convergence statistics for the probabilistic inversion of the Calhoun CZO data. ERT stands for probabilistic inversion of DC resistivity data with the empirical-Bayes-within-Gibbs approach, RS refers to empirical-Bayes-within-Gibbs inversion of the seismic refraction observations and Joint refers to the probabilistic joint inversion of the two datasets with interface constraints. The sub-scripts (1,2,3) indicate the three different chains used for each inversion routine. AR stands for the averaged acceptance rate of the chains after the burn-in and the sub-scripts I, v and r refer respectively to the interface, velocity and resistivity fields. Finally the burn-in columns enumerate the number of iterations needed for each chain to converge to the target posterior distribution [Geweke, 1992].

Chain	AR_I [%]	AR_v [%]	AR_r [%]	Burn-in$_v$ [n°of iterations]	Burn-in$_r$ [n°of iterations]
ERT_1	6.3	-	16.4	-	3.0×10^5
ERT_2	6.1	-	16.7	-	3.4×10^5
ERT_3	7.2	-	16.6	-	3.5×10^5
RS_1	9.1	16.6	-	1.5×10^5	-
RS_2	9.4	16.8	-	1.7×10^5	-
RS_3	9.0	16.9	-	1.9×10^5	-
$Joint_1$	4.8	15.1	16.9	1.9×10^5	3.8×10^5
$Joint_2$	5.2	15.5	16.8	2.0×10^5	3.5×10^5
$Joint_3$	4.6	15.1	16.7	2.1×10^5	4.0×10^5