submitted to *Geophys. J. Int.*

# Probabilistic inference of subsurface heterogeneity and interface geometry using geophysical data

G. de Pasquale[1], N. Linde[1], J. Doetsch[2], W.S. Holbrook[3]

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

[2] *Department of Earth Sciences, ETH Zürich*

[3] *Department of Geosciences, Virginia Tech*

**SUMMARY**

Geophysical methods provide remotely sensed data that are sensitive to subsurface properties and interfaces. Knowledge about discontinuities is important throughout the Earth sciences: for example, the saltwater/freshwater interface in coastal areas drive mixing processes; the temporal development of the discontinuity between frozen and unfrozen ground is indicative of permafrost development; and the regolith-bedrock interface often plays a predominant role in both landslide and critical-zone investigations. Accurate detection of subsurface boundaries and their geometry is challenging when using common inversion routines that rely on smoothness constraints that smear out any naturally-occurring interfaces. Moreover, uncertainty quantification of interface geometry based on such inversions is very difficult. In this paper, we present a probabilistic formulation and solution to the geophysical inverse problem of inferring interfaces in the presence of significant subsurface heterogeneity. We implement an empirical-Bayes-within-Gibbs formulation that separates the interface and physical property updates within a Markov chain Monte Carlo scheme. Both the interface and the physical properties of the two sub-domains are constrained to favor smooth spatial transitions and predefined property bounds. Our methodology is demonstrated on synthetic and actual surface-based electrical resistivity tomography datasets, with the aim of inferring regolith-bedrock interfaces. Even if we are unable to achieve formal convergence of the Markov chains for

all model parameters, we demonstrate that the proposed algorithm offers distinct advantages compared to manual or algorithm-based interface detection using deterministic geophysical tomograms. Moreover, we obtain more reliable estimates of bedrock resistivity and its spatial variations.

**Key words:** Inverse theory, Tomography, Probabilistic distributions, Electrical resistivity tomography (ERT)

## 1 INTRODUCTION

Near-surface geophysical datasets are inherently sensitive to physical properties of the subsurface and their spatial discontinuities. The underlying interests and reasons for mapping subsurface boundaries in the Earth sciences is discipline-dependent. For instance, knowledge about saltwater/freshwater interfaces and their evolution in coastal areas is crucial for informed groundwater management (Bear *et al.* 1999; Kim *et al.* 2006; Werner *et al.* 2013). Both aquifer overexploitation and catastrophic events (e.g., hurricanes, tsunamis) cause seawater intrusion that leads to contamination of freshwater reserves. The mixing zone between lighter freshwater and denser seawater is very difficult to study using traditional borehole monitoring data (Carrera *et al.* 2010). Geophysical methods that are sensitive to electrical conductivity are suitable to detect and study this interface and its dynamics (Nguyen *et al.* 2009; Falgas *et al.* 2009). Another interface of importance is the one between frozen and unfrozen ground (Kneisel *et al.* 2008). The evolution of permafrost impacts climate both directly by influencing heat exchange with the atmosphere and indirectly as its melting is responsible for greenhouse gas emissions (Dafflon *et al.* 2016). Geophysical detection and monitoring of this interface is most commonly achieved using seismic, electromagnetic and electrical methods (Minsley *et al.* 2012). Yet another example is landslide studies, in which near-surface geophysics plays a central role when investigating slip and failure surfaces (Jongmans & Garambois 2007) as they often correspond to the contact between regolith and bedrock, with its contrast in physical properties. Imaging with focus on the depth to the bedrock and its topography can also provide useful constraints in the study of landslide triggering mechanisms (Lanni *et al.* 2013; Reneau & Dietrich 1987; Sanchez *et al.* 2010). Moreover, bedrock mapping

is important when characterizing the critical zone (CZ): the near-surface and external layers of the Earth that include most of the life-sustaining resources on our planet (Brantley *et al.* 2007). Here, knowledge about the depth to the regolith-bedrock interface (i.e., the lower boundary of the critical zone) is necessary to model the balance between soil formation and erosion, which is important for understanding landscape evolution, sediment transport and budgets (St. Clair *et al.* 2015; Heimsath *et al.* 1997; Rempe & Dietrich 2014; Parsekian *et al.* 2015). Non-invasive geophysical investigation techniques have been increasingly applied to map regolith-bedrock interfaces and, more specifically, electrical resistivity tomography (ERT) has been commonly used because of its comparatively high resolution over the relevant depth range and the strong sensitivity of electrical conductivity to variations of relevant hydrogeological and geological subsurface properties (Chambers *et al.* 2012; Saas 2007).

Even though geophysical datasets are sensitive to such subsurface discontinuities, inferred interface locations and geometries are usually obscured by common inversion routines that seek smoothly-varying subsurface models. Geophysical inversion methods seeks to provide quantitative information about physical properties from indirect geophysical observations. It is generally an ill-posed problem, which implies that a solution based on data alone is neither stable nor unique (Kabanikhin 2008). Hence, when working with finely discretized property fields it is necessary to include a model regularization term (deterministic approach) or defining a prior probability density function (pdf) on the model parameters (Bayesian inference). In deterministic inversions, the regularization term often imposes smooth spatial variations of the Earth's subsurface by penalizing a measure of model variability (Constable *et al.* 1987; deGroot-Hedlin & Constable 1990). Because of the nature of the constraints and the non-constraining nature of the data, sharp interfaces are generally smeared out within a thick transition zone. In practice, this implies that it is up to the interpreter to identify the most likely interface geometry (Christensen 2018).

To decrease such subjectivity, Hsu *et al.* (2010) propose an automatic approach to bedrock edge detection, which analyzes 2-D smoothness-constrained deterministic inversion models obtained from an ERT dataset using image analysis. This is achieved by a gradient method (i.e., Laplacian edge detection), which searches for zero values in the second derivative of the resistiv-

ity image in both the horizontal and vertical direction. The main issue with such a methodology is the possible interference from other sources of subsurface heterogeneity that are unrelated to the large-scale interfaces of interest. To reduce the complexity of such a methodology, especially when considering 3D subsurface models, Chambers *et al.* (2012) only consider the maximum resistivity gradient in the vertical direction to locate the bedrock beneath a river terrace. Another approach that uses a guided fuzzy clustering algorithm has been proposed by Ward *et al.* (2014). Also in this case, the edge detection of bedrock interfaces is based on analysis of ERT inversion results.

Instead of imposing smoothness by model regularization, it might be more reasonable to postulate that physical properties of the subsurface are characterized by a layered structure (homogeneous layers of varying topography with distinctly different properties). Clearly, targeting the inversion towards the automatic detection of interfaces instead of producing smooth minimum-structure models and then, automatically or manually, identifying the expected boundaries from these smoothly varying images has its advantages. Within such a framework, Auken & Vest Christiansen (2004) propose a two-dimensional deterministic inversion scheme with lateral smoothness constraints and sharp vertical boundaries. Similarly, Juhojuntti & Kamm (2015) propose a method for joint inversion of seismic refraction and resistivity data using sharp-boundary models with few layers. Hence, this method produces models with laterally-varying properties within the layers, but without vertical variations within each layer. Another scenario occurs when interface locations are well known, but the fields of physical layer properties are of interest. In this case, it is possible to pre-determine unit geometries and infer physical properties and variations within each unit (Doetsch *et al.* 2012; Uhlemann *et al.* 2017; Vignoli *et al.* 2017).

In this paper, we propose a probabilistic formulation and solution to the geophysical inverse problem of inferring an interface separating two heterogeneous sub-domains. We rely on Bayes' theorem to combine the prior distribution about model parameters with a likelihood function that describes the probability of observing the collected data given proposed model parameter values and a data noise model (Sambridge & Mosegaard 2002). The resulting posterior distribution is generally not known analytically. Instead, it must be sampled numerically using, for instance, Markov chain Monte Carlo (MCMC) methods (Sambridge & Mosegaard 2002). When performing

MCMC inversion for high-dimensional distributed models, the choice of the prior pdf has a very strong influence on inversion results (Hansen *et al.* 2012, 2016; de Pasquale 2017). To favor posterior model realizations with smoothly varying properties, a possible approach is to include structure constraints (Rosas-Carbajal *et al.* 2014; Besag *et al.* 1995; Chen *et al.* 2012) within an empirical Bayes (EB) inversion framework (Casella 1985).

Here, we seek to probabilistically infer subsurface interface geometry and heterogeneities within the sub-domains that are delimited by the interface. To do so, we follow Iglesias *et al.* (2014) and implement an empirical-Bayes-within-Gibbs methodology that separates the interface and physical property updates within the MCMC scheme. More specifically, we address the particular problem of using ERT datasets to infer the regolith-bedrock interface, at a site in the Calhoun Critical Zone Observatory in South Carolina, USA. Both the interface and the physical properties of the sub-domains are constrained to favor smooth spatial variations and to honor pre-defined property bounds. We demonstrate our methodology using synthetic and actual surface-based ERT datasets. We stress that the methodology is general and that it could be used to study other types of interfaces (see examples above) and that other types of geophysical methods (e.g., seismic refraction, magnetotellurics) could be used.

We first introduce the theoretical background of the proposed empirical-Bayes-within-Gibbs inversion algorithm (Section 2), followed by the results obtained when applying it to two synthetic and a real dataset (Section 3). We then discuss our approach and findings (Section 4) before we conclude (Section 5).

## 2 METHODOLOGY

To address the challenge of probabilistic inference of subsurface interfaces and sub-domain heterogeneities, we present an inversion algorithm which combines model-structure constraints within an empirical Bayes framework with Metropolis-within-Gibbs MCMC sampling.

### 2.1 Structure-constrained empirical Bayes inversion

By combining a prior probability density function (pdf) of the model parameters ($\rho(\mathbf{m})$), summarizing our a priori information about the subsurface, and the likelihood function, which expresses the probability of the proposed model vector ($\mathbf{m}$) given the available data ($\mathbf{d}$): $\rho(\mathbf{d}|\mathbf{m}) \propto L(\mathbf{m})$ (Tarantola 2005), Bayes Theorem expresses the posterior pdf of the model parameters given the data:

$$\rho(\mathbf{m}|\mathbf{d}) = \frac{\rho(\mathbf{m})L(\mathbf{m})}{\rho(\mathbf{d})}, \tag{1}$$

where $\rho(\mathbf{d})$ only acts as a normalizing constant when the model parametrization is fixed. Analytical solutions to equation (1) are not available for non-linear problems and it is thus necessary to numerically sample from the posterior distribution. To do so, MCMC methods are often used to construct Markov chains in the model space, whose steady state distribution corresponds to the posterior distribution (Gamerman & Lopes 2006), with the sampling algorithm being completely defined by the transition probabilities. Metropolis-Hastings (M-H) is a commonly used MCMC sampler (Metropolis *et al.* 1953; Hastings 1970), which defines the transition probability from different states of the chains as being proportional to the acceptance ratio:

$$\alpha_{MH} = min\left(1, \frac{L(\mathbf{m}^{prop})\rho(\mathbf{m}^{prop})q(\mathbf{m}^{prop} \rightarrow \mathbf{m}^{curr})}{L(\mathbf{m}^{curr})\rho(\mathbf{m}^{curr})q(\mathbf{m}^{curr} \rightarrow \mathbf{m}^{prop})}\right), \tag{2}$$

where $q$ is the proposal pdf describing the probability to generate a given model perturbation at each proposal step, and the superscripts *prop* and *curr* refer, respectively, to the proposed and the current state of the Markov chain. Acceptance ratios used in MCMC codes are typically reformulated using a logarithmic formulation to avoid floating-point under- and overflow. Most MCMC applications in geophysics rely on a symmetric proposal pdf, $q(\mathbf{m}^{prop} \rightarrow \mathbf{m}^{curr}) = q(\mathbf{m}^{curr} \rightarrow \mathbf{m}^{prop})$, which leads to the simpler Metropolis acceptance ratio:

$$\alpha_{M} = min\left(1, \frac{L(\mathbf{m}^{prop})\rho(\mathbf{m}^{prop})}{L(\mathbf{m}^{curr})\rho(\mathbf{m}^{curr})}\right), \tag{3}$$

for which there is no evaluation of the proposal pdfs at each step of the chain.

When the prior knowledge about the subsurface is limited, it is common to consider uncorrelated parameters with uniform probability over a given parameter range. This choice of priors

leads to posterior model realizations that are too spatially variable to be geologically realistic when applied to high-dimensional distributed models (Rosas-Carbajal *et al.* 2014; Hansen *et al.* 2012, 2016; de Pasquale & Linde 2017). An alternative is to also constrain global model structure by penalizing model variability. The appropriate weight given to these constraints can be obtained through an empirical Bayes approach.

Empirical Bayes (EB) inference (Casella 1985) uses hyper parameters describing the prior distribution (e.g., its standard deviation, integral scales). The hyper parameters are then estimated together with the regular model parameters during the inversion process. When using MCMC, the hyper parameters require both prior distributions and proposal pdfs. For the pdf describing model structure used herein, the hyper-parameter $\lambda$ defines the mean deviation of the exponential model:

$$c(S(\mathbf{m}), \lambda) = (2\lambda)^{-Q} \exp\left(-\frac{||S(\mathbf{m})||_1}{\lambda}\right), \tag{4}$$

where $Q$ indicates the total rank of the model structure operator and the measure of model structure $S(\mathbf{m})$ is computed with an $l_1$-norm (a corresponding formulation is used to define interface roughness) as further explained in section 2.3. Similar to the model regularization weights used in deterministic inversion, a high $\lambda$ strongly penalizes model structure. Considering symmetric model proposal pdfs for the model parameters and the hyper-parameter, the acceptance ratio in case of structure-constrained empirical Bayes MCMC inversion is:

$$\alpha_{EB} = min\left(1, \frac{L(\mathbf{m}^{prop})\rho(\mathbf{m}^{prop})\rho(\lambda^{prop})c(S(\mathbf{m}^{prop}), \lambda^{prop})}{L(\mathbf{m}^{curr})\rho(\mathbf{m}^{curr})\rho(\lambda^{curr})c(S(\mathbf{m}^{curr}), \lambda^{curr})}\right), \tag{5}$$

where $\rho(\lambda)$ is the prior pdf of the hyper parameter.

In this study, we considered $\rho(\lambda)$ and $\rho(\mathbf{m})$ to have log-uniform and uncorrelated uniform pdfs, respectively, such that the acceptance ratio simplifies to the ratio involving likelihood and constraint functions only.

## 2.2 Empirical-Bayes-within-Gibbs

We conceptualize the subsurface as being composed of two overlapping domains, $\mathbf{m}_{CZ}$ and $\mathbf{m}_b$, that are separated by an interface, $\mathbf{I}$. We first draw two uncorrelated physical property fields that are parametrized with regular grids that cover the whole investigated area (Fig. 1a). In the examples

that follow, we used $50 \times 10$ grid with cell sizes of $8 \times 8$ meters. Herein, the physical properties considered are the logarithm of electrical resistivity $r$ [$\Omega$m], that is discretized throughout the sub-domains representing the critical zone (CZ, $\mathbf{m}_{CZ}$) and the bedrock ($\mathbf{m}_b$). The forward operator is defined on an unstructured mesh that covers the investigated area (Fig. 1b) and extends on the sides and below to account for boundary conditions. In the examples, the unstructured mesh used for forward simulations is discretized with 1578 cells. We linearly interpolate the sampled electrical resistivity values into the unstructured mesh (Fig. 1c). In parallel, we draw an interface from the corresponding prior pdf, $\rho(\mathbf{I})$. In 2D, the interface defines a line delineated by a set of connected nodes within the mesh (Fig. 1d). The actual resistivity field ($\mathbf{m}$) used for forward simulations is built by mapping each of the interpolated physical property values into the appropriate sub-domain defined by the interface (Fig. 1e):

$$\mathbf{m} = G(\mathbf{m}_{CZ}, \mathbf{m}_b, \mathbf{I}), \tag{6}$$

where $G$ is the mapping operator.

Iglesias *et al.* (2014) introduced a Metropolis-within-Gibbs method in which the model proposals are symmetric and the interface and physical properties are updated alternately within the Gibbs framework (Geman & Geman 1984). In our implementation, we additionally constrain the interface and the resistivity fields to favor smoothness, resulting in a slightly modified empirical-Bayes-within-Gibbs algorithm in which we alternately propose an update to the interface or to the resistivity field of the domains (Fig. 2 shows a flowchart of our algorithm).

The interface proposal updates are implemented as follows. Initially, one vertex on the current interface is chosen at random and removed from the interface definition. Then, one of its two adjacent vertices in the current interface is picked at random. When this vertex (anchor) is located on the left, one of its adjacent vertices in the mesh is drawn at random under the constraint that it is located to the right and that it is not the previously removed vertex. From this point, the shortest path is sought to a vertex belonging to the current interface under the constraint that all horizontal increments in the path are to the right. The vertices of this path are now part of the proposed interface geometry. When the anchor is located to the right, the model proposal procedure

is adapted in a straightforward manner. In this way, we ensure symmetry in our model proposals and avoid a possible situation with the bedrock overlying itself. The probability of accepting the proposed interface is given by:

$$\alpha_I = min\left[1, \frac{L(\mathbf{m}_{CZ}^{curr}, \mathbf{m}_b^{curr}, \mathbf{I}^{prop}) c(R(\mathbf{I}^{prop}), \lambda_{\mathbf{I}}^{prop})}{L(\mathbf{m}_{CZ}^{curr}, \mathbf{m}_b^{curr}, \mathbf{I}^{curr}) c(R(\mathbf{I}^{curr}), \lambda_{\mathbf{I}}^{curr})}\right], \tag{7}$$

where $R(\mathbf{I})$ quantifies the interface roughness and the likelihood ratio is computed using the model obtained after updating the interface ($\mathbf{m}_{CZ}^{curr}, \mathbf{m}_b^{curr}, \mathbf{I}^{prop}$) and the model accepted in the previous step of the chain ($\mathbf{m}_{CZ}^{curr}, \mathbf{m}_b^{curr}, \mathbf{I}^{curr}$).

The electrical resistivities are described using priors defined on different ranges: $\rho(\mathbf{m}_{CZ}) \neq \rho(\mathbf{m}_b)$, as bedrock is generally more resistive than the overlying CZ. To update model properties, we choose a certain fraction of model parameters to update, randomly disposed within the domain, and sample the new values from a Gaussian pdf centered on the previous value. Moreover, with the aim of augmenting the posterior model space exploration, we alternate the choice of step size for the proposal distributions by switching between applying a randomly drawn standard deviation value from a Gaussian pdf and applying it uniformly over the whole domain or scaling it inversely to the square root of the cumulative sensitivity of the investigated subsurface. Model proposals that use the same standard deviation throughout the model domain help to converge to the target misfit, while the proposals based on the cumulative sensitivities allows for improved exploration in regions of low sensitivity. The cumulative sensitivities are here obtained from the model result of a deterministic inversion, but the outcomes are expected to be very similar if using the sensitivities obtained from a uniform model. Indeed, such a scaling of the standard deviations does not impact the symmetry and stationarity of the proposal distribution throughout the MCMC procedure, hence, it is a theoretically-valid model proposal. The model proposals are implemented for both domains at the same time and they can involve parameters that are not mapped into the resistivity field used for the forward computation (e.g., a parameter describing CZ resistivity at the bottom of the model domain). The corresponding acceptance ratio for the physical properties update is:

$$\alpha_P = min\left[1, \frac{L(\mathbf{m}_{CZ}^{prop}, \mathbf{m}_b^{prop}, \mathbf{I}^{curr}) c(S(\mathbf{m}_{CZ}^{prop}), \lambda_{CZ}^{prop}) c(S(\mathbf{m}_b^{prop}), \lambda_b^{prop})}{L(\mathbf{m}_{CZ}^{curr}, \mathbf{m}_b^{curr}, \mathbf{I}^{curr}) c(S(\mathbf{m}_{CZ}^{curr}), \lambda_{CZ}^{curr}) c(S(\mathbf{m}_b^{curr}), \lambda_b^{curr}))}\right], \tag{8}$$

where the model structure ($S(\mathbf{m})$) is computed for both domains (i.e., not only for the sub-domains found below or above the interface). We use different constraint functions for CZ and the bedrock, and sample two different hyper-parameters ($\lambda_{CZ}$ and $\lambda_b$). The likelihood ratio is computed between the model obtained by updating the two resistivity fields ($\mathbf{m}_{CZ}^{prop}, \mathbf{m}_b^{prop}, \mathbf{I}^{curr}$) and the current model of the chain ($\mathbf{m}_{CZ}^{curr}, \mathbf{m}_b^{curr}, \mathbf{I}^{curr}$).

### 2.3 Measures of model structure and interface roughness

As mentioned above, we consider two different measures of model structure: one quantifying the spatial variability of the subsurface resistivity field (i.e., $S(\mathbf{m})$ in eq. 8) and another quantifying the roughness of the interface (i.e., $R(\mathbf{I})$ in eq. 7). In the first case, we consider a measure of model structure that is widely used in deterministic inversion (Menke 1989), namely $l_1$-norms of model roughness, and we apply it to the inversion parameter grid (i.e., the regular grid parametrization represented in Fig. 1a). We follow de Pasquale & Linde (2017) by describing each grid element (i.e., model parameter) with two indices: $m_{i,j}$. Considering the grid analogously to a matrix, the measure of electrical property variability can be expressed as:

$$S(\mathbf{m}) = \sum_{|i-l|\leq 1} \sum_{|j-k|\leq 1} |m_{i,j} - m_{l,k}| \qquad |i-l| + |j-k| = 1. \tag{9}$$

To quantify the interface roughness, we use a corresponding $l_1$ metric of the differences between z-coordinates (e.g., depth) of the interface vertex, according to the following equation:

$$R(\mathbf{I}) = \sum_{l=1}^{n-1} |z_{l+1} - z_{l-1}|, \tag{10}$$

where $l = 0 \ldots n$ refers to the nodes intersected by the interface $\mathbf{I}$, with index $l$ growing with the profile length (i.e., from left to right in Fig. 1d). Initial tests (not shown) suggested superior inversion results when considering a central difference measure compared with a forward difference measure.

### 2.4 ERT forward modeling

In the following test cases, we consider 2D electrical resistivity tomography datasets. Electrical resistivity tomography is widely used to determine the spatial distribution of electrical resistivity.

In practice, apparent resistivity data are obtained using a large number of resistance measurements made from spatially-distributed four-point electrode configurations (Dahlin 1996). In order to simulate the ERT forward response needed to compute the likelihood of each model realization within the Markov chain, we used the Boundless Electrical Resistivity Tomography (BERT) library, which is an open source software package widely used for deterministic inversion (Rücker *et al.* 2006; Günther *et al.* 2006). Using BERT, we can compute the 2.5D forward response by relying on a finite element scheme implemented on unstructured meshes (Si 2015). Using unstructured meshes allow for a more efficient local refinement (e.g., close to electrodes) and are ideally suited to account for surface topography and internal boundaries.

## 3 RESULTS

To evaluate the performance of our proposed algorithm, we consider a 2D surface-based ERT dataset acquired above a regolith-bedrock interface. More specifically, we refer to data from the Calhoun Critical Zone Observatory (South Carolina, USA) presented 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. The ERT survey was acquired using an Advanced Geosciences Inc's (AGI) SuperSting instrument with 56 electrodes spaced 5 m apart with a dipole-dipole electrode configuration. The profile length of 400 m was obtained by using one roll-along in which 50 % of the electrodes stayed in place. Consequently, St. Clair *et al.* (2015) used 84 different electrode positions. To save computational time and since our interest is in the deeper CZ structure, we only consider every second electrode position and we removed the configurations with a maximal electrode separation of 30 m. Moreover, the raw data were filtered by St. Clair *et al.* (2015) to eliminate negative apparent resistivity values. This resulted in 42 electrodes and 645 data being used. A noise description based on a relative Gaussian error of 3.7 % was used in order to reproduce the same data misfit as the inversion results by St. Clair *et al.* (2015).

### 3.1 Synthetic test cases

To test the methodology in a controlled environment, we first consider two synthetic examples with the same electrode configuration and surface topography as for the real dataset. Both examples present the same interface geometry, but the first example features homogeneous sub-domains and the second has heterogeneities within the layers.

To demonstrate that the empirical-Bayes-within-Gibbs algorithm can sample the target interface, we test the method on the model shown in Fig. 3(a). In this case, the interface divides the domain in two homogeneous domains: an upper layer with resistivity of 1800 Ωm and a lower layer with resistivity of 15000 Ωm. Fig. 3(b) represents the interpolation (through the nearest neighbor method) of the input resistivity model on the mesh used for the forward computations within the MCMC chains. Note that this mesh is different from the mesh used to generate the synthetic observations. The best interface we can sample is represented by a black line in Figs. 3(c) and (e). We contaminate the simulated observations with 3.7 % uncorrelated Gaussian noise. The mean model sampled by the empirical-Bayes-within-Gibbs chain in Fig. 3(c) provides a very close approximation of the interface; moreover, the posterior realization corresponding to the maximum likelihood (Fig. 3d) exactly reproduces the interface of the interpolated target model. We contrast these results with the deterministic inversion result, in which the interface is obtained from the maximum vertical gradient method (i.e., the blue line in Fig. 3e ) using the approach described in Chambers *et al.* (2012). From these results, it is evident that the empirical-Bayes-within-Gibbs approach is more successful in sampling the target interface. To quantify this, we compute the average distance between the target interface ($I_{target}$) and the sampled (in case of empirical-Bayes-within-Gibbs) or inferred one (in case of deterministic inversion result):

$$\bar{D} = \frac{\int_{x_{min}}^{x_{max}} |I_{target} - I_{sampled/inferred}|}{x_{max} - x_{min}}, \tag{11}$$

where $x_{min}$ and $x_{max}$ are the horizontal limits of the model domain. For the deterministic inversion this distance is 2.9 m, while when evaluating the measure on a sample of posterior realizations,

we find that the empirical-Bayes-within-Gibbs algorithm yields an average distance of $0.4 \pm 0.1$ m only.

After verifying that the algorithm samples the target interface for this simple example, we introduce significant heterogeneities within the sub-domains. In Fig. 4(a), we show the corresponding resistivity model used to generate the new synthetic dataset together with the electrode positions. The underlying mesh is constructed to exactly accommodate the defined interface, while the two heterogeneous sub-domains are obtained by stationary multivariate Gaussian process generations through circulant embedding of the covariance matrix (Dietrich & Newsam 1997). Fig. 4(b) represents the interpolation (through nearest neighbor method) of the input resistivity model on the mesh used for the forward computations within the MCMC chains. This image gives an idea of the best possible representation of the interface that can be sampled within the empirical-Bayes-within-Gibbs inversion. As before, the simulated data were subsequently contaminated with 3.7 % uncorrelated Gaussian noise (Fig. 4c).

Within our inversion routine, we used intervals for the electrical resistivity fields according to typical ranges of CZ and bedrock in Fig. 4(b): $\mathbf{m}_{CZ} \in [100, 3000]$ $\Omega$m and $\mathbf{m}_b \in [5000, 30000]$ $\Omega$m. As described in Section 2.2, the resistivity fields are updated by randomly choosing 10 % of the corresponding model parameters (i.e., varying 10 % of the resistivity values in the regular grids in Fig. 1a). At each $\mathbf{m}_{CZ}$ and $\mathbf{m}_b$ model proposal step, the model parameters are updated using a Gaussian distribution centered on the current model with a standard deviation between 0.002 and 0.02: $\sigma \sim U(0.002, 0.02)$. These values were chosen to ensure an acceptance rate close to 25 % and to accommodate both smaller and larger model updates. We alternate between using a uniform value as standard deviation for all the model parameter updates and reweighing them inversely to the square root of the cumulative sensitivities (Fig. 4d). Based on extensive testing, we find that alternating model updates in this way enables improved exploration in regions of low sensitivity while still ensuring that the residuals of the model predictions have the same chi-square mean as the assumed noise term (i.e., that we reach the target misfit).

The prior on the interface is defined by the way the interface can be constructed within the unstructured mesh used for computing the model response: nodes defining an interface have to be

connected between each other through edges of the forward mesh and the horizontal increment must be either always positive or always negative within one realization (i.e., as in Fig. 1b). For the prior definition of model structure or interface roughness weights used to constrain the model complexity (i.e., the hyper-parameters $\lambda_{CZ}$, $\lambda_b$ and $\lambda_I$ introduced in Sections 2.1 and 2.2), we follow Rosas-Carbajal *et al.* (2014) and use log-uniform priors with ranges between one quarter and six times the values found when eq. 4 is maximized for the true model. In the case of interface roughness, the maximization is done considering the value obtained by the interface of the model depicted in Fig. 4(b). For CZ and bedrock, we use the same value that was obtained by considering the structure of the full model (i.e., computed from the model represented in Fig. 4c). For the hyper-parameters, we consider Gaussian proposal distributions with a constant standard deviation of $\sigma_\lambda = 0.1$.

To sample the posterior space, we run three independent chains for $10^6$ iterations with each iteration taking $\approx 0.6$ seconds. The starting model of the three chains are presented in Figs. 5(a), (d) and (g), where the logarithmic resistivity fields are sampled with uniform uncorrelated prior probability. The corresponding model predictions have weighted root mean squared errors (*WRMSE*) of 28, 34 and 35, which is much higher than the target value of 1 on which the posterior distribution is centered. That is, the misfits of the initial models are some 30 times larger than the posterior realizations. The last posterior model realizations (Figs. 5b, e, h), as well as the maximum likelihood models (Figs. 5c, f, i) present resistivity fields and bedrock topography that are largely consistent with each other and with the underlying true model (Fig. 4b).

The "posterior" distributions are constructed based on the sampled realizations within the stationary part of the chains (i.e., after burn-in when the chains start to sample proportionally to the target distribution). To evaluate the burn-in length for each chain, we rely on the Geweke diagnostic (Geweke 1992). 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. Here, this diagnostic results in a burn-in of $3 \times 10^5$, $2.5 \times 10^5$ and $4 \times 10^5$ itera-

tions for the first, second and third chain, respectively. Figs. 6(a) and (b) show the sampled pdfs of model structure for CZ and bedrock resistivity fields, respectively, while Fig. 6(c) represents the interface roughness. The marginal distribution of the hyper-parameter used as structure-constraint weights are depicted in Figs. 6(d)-(f), together with the optimal values for the true model. The mean acceptance rate for the interface updates (Fig. 6g) is lower (average value of 8 %) than for the resistivity field updates (Fig. 6h; average value of 25 %). Finally, Fig. 6(i) represents the log-likelihood evolutions for the three chains. The potential scale reduction factor (Gelman & Rubin 1992) is commonly used to assess if a sufficient number of posterior samples have been obtained in order to adequately describe the posterior distribution. For each model parameter, it compares the between-chain and the within-chain variance. In practice, it is generally agreed that a value of 1.2 is sufficient to declare convergence. Considering the last $5 \times 10^5$ iterations, the potential scale reduction factor is found to be below 1.2 for only 16 % of the CZ parameters and 45 % of the bedrock parameters. Therefore, these results indicate that we did not sufficiently sample the posterior pdf.

Nevertheless, we have seen that the different chains produce similar posterior model realizations (Fig. 5). The mean values (Figs. 7a, d and g) and normalized standard deviations of electrical resistivity (i.e., divided by the mean resistivity values; Figs. 7b, e and h) are comparable for all the chains. This strong similarity is also seen in the probability maps of the interface locations (Figs. 7c, f and i), and in the vertical resistivity profiles (Figs. 7l, m and o). For all chains we see that the empirical-Bayes-within-Gibbs samples the underlying interface well, except for the left side of the domain (first 50 m of the profile length) and around 370 m along the profile, where the inferred bedrock interface is too shallow. Moreover, all the three chains show the highest normalized standard deviation in the vicinity of the inferred interface. In this region, the interface is varying throughout the MCMC chain and the same cell is alternately associated with the CZ (more conductive) or bedrock (more resistive).

In order to evaluate the quality of the sampled posterior model realizations, we computed the model discrepancy (Fig. 8b):

$$w_r = \mathbf{m}_M - \mathbf{m}, \tag{12}$$

where $\mathbf{m}_M$ is the vector of the $\log_{10}$ resistivity values of the true model (Fig. 6b) and $\mathbf{m}$ is the mean of the posterior model realizations (Fig. 8a) for the three Empirical-Bayes-within-Gibbs chains. For comparison purposes, we also present the inversion model result obtained by a traditional $l_1$-norm smoothness-constrained deterministic inversion that is fitted to the same error level (Fig. 8c) and the corresponding model discrepancy (Fig. 8d). We find that the MCMC inversion strongly overestimates the resistivity of the upper 40 m at the beginning of the profile, while the deterministic inversion consistently underestimates the resistivity of the bedrock, this is particularly visible in the vertical resistivity profiles represented in Figs. 8(e)-(h). The mean absolute discrepancy in Fig. 8(b) is lower (0.19) than for the deterministic inversion (0.32 in Fig. 8d). Fig. 8(c) shows also the interface obtained from the deterministic inversion using the maximum vertical gradient method. Also in this case, it is seen that the estimated interface is the worst on the left side of the parameter domain. The distances to the true interface (i.e., eq. 11) are comparable for the two types of inversions: for the deterministic inversion the average distance is 3.9 m, while the distance for the different empirical-Bayes-within-Gibbs chains are 3.6 $\pm 0.1$ m, 4.5 $\pm 0.2$ m and 4.1 $\pm 0.1$ m. The error in the MCMC inversion is almost entirely due to the poor performance in the first 50-100 m of the profile. Note that the problematic zones are on the sides of the domain where the ERT data do not constrain the model properties well, while the agreement is good in the central part where the data are more constraining.

### 3.2 Field data

As already discussed above, we filter the ERT dataset from the Calhoun Critical Zone Observatory (CZO) in South Carolina (USA). Indeed, when considering all 84 electrodes, the chains were unable to converge to the target data misfit. We attribute this to significant small-scale heterogeneity in the near-surface below our model discretization size (see Fig. 1a). We could have refined the inversion grid, but we chose instead to decrease the data sensitivity to small-scale near-surface variability by not considering neighboring electrodes and, therefore, removing the shallowest pseudo depth level from the data. This choice is motivated by our focus on the bedrock interface and that a model refinement would have led to unnecessarily high computational times.

In agreement with the synthetic test case, we sampled the electrical resistivity, model structure and the interface roughness weights from log-uniform distributions with the structural constraints expressed in Eq. (4). Here, the prior range of the hyper parameters are defined broadly enough to avoid boundary effects (see Rosas-Carbajal *et al.*, 2014). To establish the prior range on resistivities, we first determined a global resistivity range (i.e., over the whole subsurface), which was obtained from the deterministic inversion results when considering the full dataset (i.e., considering the full 84-sensor dataset). Our target interface is the one between regolith (weathered/fractured bedrock) and unweathered bedrock, which is clearly seen at 20-40 m depth in seismic and resistivity data presented by St. Clair *et al.* (2015). In those results, unweathered bedrock occurs where seismic velocity increases to more than 4 km/s and resistivity increases to more than 4000 $\Omega$m. We had to run different tests to establish an upper limit for the CZ resistivity range and a lower limit for the bedrock. The final ranges of $\mathbf{m}_{CZ} \in [80, 10000]$ $\Omega$m and $\mathbf{m}_b \in [2000, 30000]$ $\Omega$m, were chosen to ensure the convergence to the target data misfit. The physical parameter and interface updates, as well as the priors and proposal pdfs for the hyper-parameters, are implemented in the same way as for the synthetic test case with heterogeneous sub-domains. We initialize the chains with random models (Figs. 9a, d and g); the corresponding *WRMSE* of these models are 28, 25 and 34. By comparing the three chains, it is seen that the last posterior model realizations (Figs. 9b, e, h) and the maximum likelihood models (Figs. 9c, f, i) present similar resistivity fields and bedrock topography. Particularly, all the realizations show a weak zone within the bedrock at around 300 m along the profile.

After MCMC inversion, we find that the inferred model structure for the CZ (Fig. 10a), the bedrock (Fig. 10b) and the interface (Fig. 10c) are overall similar for the three chains, as are the hyper-parameters (Figs. 10d-f). In agreement with the synthetic results, the acceptance rate for the interface updates (Fig. 10g) is lower (average value of 6 %) than for the resistivity field updates (Fig. 10h; average value of 17 %). According to the Geweke diagnostic, the burn-in is $5 \times 10^5$, $4.5 \times 10^5$ and $3.5 \times 10^5$ iterations for the first, second and third chain, respectively. As for the synthetic test case, we could declare formal convergence for only a percentage of the model parameters: 14 % for the CZ and 23 % for the bedrock. This implies that even if we obtain a set of

models that are able to explain the observations well, the chains do not fully explore the posterior distribution.

Inversion results in terms of mean, normalized standard deviation of resistivity and bedrock probability map of the "posterior" model realizations are shown in Fig. 11. The results show a very similar behavior among the different chains, especially in the upper 40 meters (where the resistivity data are able to constrain the model parameters). This is also seen in the vertical resistivity profiles in Figs. 11(l)-(o), where the consistency in the inferred interface locations is manifested by the jumps in resistivity. For comparison purposes, we also run the empirical Bayes framework with three chains for which we did not consider the subdivision of the subsurface. That is, we infer one log-resistivity field using a prior pdf that spans the range of the previously defined sub-domains. Similarly to previous results, we start to sample the posterior probability distribution after $10^5$ iterations and we are again unable to fully sample the posterior distribution. The corresponding approximate posterior results in terms of mean, normalized standard deviation and vertical resistivity profiles are represented in Fig. 12. Finally, we compared these results to the ones obtained by smoothness-constrained deterministic inversion for the same data misfit (Fig. 13c). For both inversions, we inferred an interface using the vertical gradient method (i.e., the black lines in Fig. 13c and d).

## 4 DISCUSSION

We have presented the first inversion approach that uses ERT data to probabilistically infer interface properties in the presence of heterogeneous sub-domains. The results are promising, but there are also certain limitations. Initially, we attempted to address the inverse problem using the more general formulation presented by de Pasquale & Linde (2017), but we were unable (despite significant testing) to obtain stable results. We attribute this to the wider prior ranges in the present work compared to the successful case-studies considered by de Pasquale & Linde (2017). Both Rosas-Carbajal *et al.* (2014) and de Pasquale & Linde (2017) have highlighted that the type of structure-constrained inversion routine implemented herein tends to favor model realizations with too little complexity (i.e., too smooth models), especially where the data are weakly sensitive.

The empirical-Bayes-within-Gibbs method successfully sampled the modeled interface in the synthetic example with homogeneous properties (Fig. 3). When we consider the synthetic test case with heterogeneous sub-domains (Fig. 4), however, the inversion results highlight that the target interface is not part of the sampled interface locations along the first 50 m and around 370 m of the profile length (see Figs. 7a, d, g and 8a). The true model in Fig. 4(a) presents a strong resistor (40 m depth and 50 m along the profile line) which is situated below a conductive area. At 370 m, the true model has a high resistivity bedrock at the bottom, which is overlain by a conductive deep-CZ anomaly and a higher resistivity feature at the surface. Due to the inherent equivalence problem of DC resistivity data (Koefoed 1979), the resulting forward response can, thus, also be explained by the posterior realizations we obtained (i.e., a large area with intermediate resistivities for both cases). Such results are favored in the empirical-Bayes inversion because they are less complex (in terms of the variability within the two sub-domains) than the underlying true model. These inconsistencies persist even when using homogeneous starting models. Moreover, Fig. 8(c) shows that the interface inferred using the maximum gradient method on the deterministic inversion model is shallower than the target one in the beginning of the profile. This inefficiency of the different inversion routines stresses the inability of the ERT data to resolve this area. Nevertheless, the model results are satisfactory in the central part of the profile and remain an advancement compared with smoothness-constrained deterministic inversion results (see Fig. 8). Perhaps the most dramatic improvement is found in the bedrock domain, whose resistivity values are always severely under-estimated by the deterministic inversion.

Even if the posterior model realizations are similar across chains, we are unable to declare convergence of the MCMC chains for all the model parameters (Gelman & Rubin 1992); 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. By combining the sampled posterior realizations of the three chains, we partially reduce this issue. Inversion for sharp boundaries between subsurface physical properties is hard as it makes the likelihood surface highly irregular (Tavassoli *et al.* 2005), especially for high-dimensional problems and large data sets with high signal-to-noise ratios. More advanced MCMC algorithms (e.g., parallel tempering) could improve

the situation (Laloy *et al.* 2016). We have found that scaling of the model proposal updates using model sensitivities improve exploration of the posterior distribution. Adaptive Metropolis (Haario *et al.* 2001) and its variants (e.g., Laloy & Vrugt 2012), in which an appropriate proposal distribution is determined for each model parameter might further improve exploration. We also suspect that some of the problems encountered in this study are further enhanced by the fact that ERT data have no (or very limited) inherent sensitivity to the depth and thickness of subsurface layers (Parker 1984; Oldenburg & Li 1999). Another possibility would be to address this problem with a transdimensional inversion algorithm (Sambridge *et al.* 2006; Belhadj *et al.* 2018), in which the complexity of the interface geometry and the physical property fields are determined by treating the model dimension of sub-domains and the number of interface locations as unknowns.

The results obtained for the Calhoun field data are summarized in Fig. 13, where we show the models obtained for the different inversion routines considered. Regardless of the inversion approach, we find that the upper part of the domain (first 10-20 meters of depth) is well defined. However, the inferred bedrock interface and bedrock properties are highly dependent on the inversion approach. As expected, we find that the smoothness-constrained deterministic inversion result displays the least features (Fig. 13c) and a smaller range of resistivity values: $\mathbf{m} \in [150, 7000]\ \Omega$m compared with the probabilistic inversion results (Figs. 13a and b) with $\mathbf{m} \in [100, 25000]\ \Omega$m. The interfaces we inferred through the maximum vertical gradient method is represented with a black line in Fig. 13(c). When considering the subsurface as a single domain, the probabilistic inversion provides large fluctuations in resistivity values (Fig. 13b) including geologically-unrealistic conductors at depth. The inferred interface (black line in Fig. 13b) shows unrealistic depth variations especially at the far left end of the domain. In contrast, the proposed empirical-Bayes-within-Gibbs algorithm ensures that no conductors are found at depth (Fig. 13a), while it still allows for a significant variability within the CZ and bedrock sub-domains. Moreover, the interface is part of the model parameterization and is, hence, clearly described in terms of its depth, geometry and uncertainties. We note that it is only this inversion that clearly identify a deepening of the inferred bedrock interface at the two topographic lows. Note further that the small-scale irregularity of the inferred interface in Fig. 13(a) is a consequence of the mesh discretization. It could be removed in a

post-processing step by smoothing the interface over a length scale corresponding to twice the average node separation or by using a finer mesh. The vertical resistivity profiles represented in Figs. 13(d)-(g) underline the different characteristics of the inversion results: the deterministic inversion results are generally smoother and they present a smaller resistivity range, while the probabilistic inversion with single subsurface domain show larger fluctuations. Finally, the empirical-Bayes-within-Gibbs vertical profiles clearly display the resistivity jump at the interface location and the widest resistivity range.

Our results are now compared with those of St. Clair *et al.* (2015) regarding the depth and shape of the regolith-bedrock transition. Those authors located the transition between CZ and bedrock at the site between 10 and 40 meters depth, with a shape that mirrors topography (deep under ridges and shallow under valleys). This pattern, which has been attributed to the effect of compressive tectonic stress on fracture opening (St. Clair *et al.* 2015; Moon *et al.* 2017) is most apparent in seismic velocity (Fig. 4 of St. Clair *et al.* 2015). Our results indicate similar depths, but rather interface patterns that follows the topography (Figs. 11 a, d, and g). For instance, all the inversion results suggest the presence of a low-resistivity zone in the valley at 300 m along the profile line ($\mathbf{m} \approx 2000\ \Omega$m). Extensive testing with many alternative prior resistivity ranges and starting models illustrate that such a zone is needed if the data are to be fitted with a realistic-looking model. A possible explanation for this subsurface resistivity pattern is the presence of deep fracture zones or faults underlying the stream valleys. To better understand this difference in behavior, we suggest that joint inversion of the seismic and ERT data would be of great interest.

## 5 CONCLUSIONS AND OUTLOOK

Determining subsurface interface geometries together with reliable uncertainty quantification is important in various Earth science settings. This objective is generally not achievable when using deterministic smoothness-constrained inversions or when inferring for interface locations while ignoring subsurface heterogeneities within layers. In this work, we introduce an empirical-Bayes-within-Gibbs MCMC inversion algorithm that explicitly parameterize and infer both interface geometry and spatial heterogeneity of physical properties. Our synthetic and field-based results

consider 2D surface-based ERT data aiming at inferring the regolith-bedrock interface, but the methodology is general and, since it is independent of the forward solver, it can be extended to other observation types (e.g., magnetotellurics or seismic refraction), 3D datasets and parametrization choices. For a synthetic heterogeneous test case, we find that the interface location is well-resolved in the central part of the profile, but less so on the sides where the data are less constraining. The introduction of the interface in the inversion leads to a dramatic improvement in the estimations of the bedrock properties when compared with smoothness-constrained deterministic inversions and MCMC-based inversions without an explicit interface. Nevertheless, longer MCMC chains or adaptations of more advanced sampling methods (e.g., parallel tempering) is needed to fully explore the posterior distribution. For the field example at the Calhoun Critical Zone Observatory, our new method suggested a significant deepening of the bedrock geometry coinciding with topographic lows. The time needed for each iteration is mainly determined by the forward operator (in the case of 2D ERT, this time scale linearly with the number of current electrodes) and the number of nodes in the unstructured forward mesh, which affects the interface update performance (also in this case, the time scales linearly with the number of nodes). A logical extension of this work is to perform joint inversion of ERT and seismic refraction data. This would help to better resolve the regolith-bedrock interface and it would overcome some of the equivalence issues that are inherent with ERT datasets. The model coupling could be achieved by considering a common interface geometry and possibly also parameter correlation of the physical property fields (e.g., to enforce that highly resistive bedrock is likely to correspond to regions of high seismic velocity).

## ACKNOWLEDGMENTS

This research was supported by the Swiss National Science Foundation under grant 200021-155924. We are grateful for the helpful comments and remarks provided by Jochen Kamm, Thomas Mejer Hansen and two anonymous reviewers.

## REFERENCES

Auken, E. & Christiansen, A. V., 2004. Layered and laterally constrained 2D inversion of resistivity data. *Geophysics* **69**(3), 752-761.

Bear, J., Cheng, A.H.-D., Sorek, S., Ouazar, D.& Herrera, I., 1999. Seawater Intrusion in Coastal Aquifers-Concepts, Methods and Practices. *Kluwer Academic Publisher* 625 p.

Belhadj, J., Romary, T., Gesret, A., Noble, M. & Figliuzzi, B., 2018. New parametrizations for Bayesian seismic tomography. *Inverse Problems* **34**, 065007(33pp).

Besag, J., Green, O., Higdon, D. & Mengerson, K., 1995. Bayesian computation and stochastic systems. *Statistical Science* **69**, 3-41.

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

Carrera, J., Hidalgo, J.J. & Slooten, L.J., 2010. Computational and conceptual issues in the calibration of seawater intrusion models. *Hydrogeology Journal* **18**, 131-145.

Casella, G., 1985. An introduction to empirical Bayes data analysis. *Am. Statisticians* **39**(2), 83-87.

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.

Chen, J., Hoversten, G.M., Key, K., Nordquist, G. & Cumming, W., 2012. Stochastic inversion of magnetotelluric data using sharp boundary parametrization and application to geothermal site. *Geophysics* **77**(4), E256-E279.

Christensen, N.B., 2018. Interpretation attributes derived from airborne electromagnetic inversion models using the continuous wavelet transform. *Near Surface Geophysics* **16**(6), 665-678.

Constable, S.C, Parker, R.L. & Constable, C.G., 1987. Occam's inversion: A practical algorithm for generating smooth model from electromagnetic sounding data. *Geophysics* **52**(3), 289-300.

deGroot-Hedlin, C. & Constable, S., 1990. Occam's inversion to generate smooth, two-dimensional models from magnetotelluric data. *Geophysics* **55**, 1613-1624.

Dafflon, B., Hubbard, S., Ulrich, C., Peterson, J., Wu, Y., Wainwright, H. & Kneafsey, T.J., 2016. Geophysical estimation of shallow permafrost distribution and properties in an ice-wedge polygon-dominated Arctic tundra region. *Geophysics* **18**(1), WA247-WA263.

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

de Pasquale, G., 2017. Changing the prior model description in Bayesian inversion of hydrogeophysics dataset. *Groundwater* **55**(5), 1342-1358.

de Pasquale, G. & Linde, N., 2017. On structure-based priors in Bayesian geophysical inversion. *Geophysical Journal International* **208**(3), 651-655.

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.

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

Falgas, E., Ledo, J., Marcuello, A. & Queralt, P., 2009. Monitoring freshwater-seawater interface dynamics with audiomagnetotelluric data. *Near Surface Geophysics* **7**(5-6), 391-399.

Gamerman, D. & Lopes, H.F., 2006. *Markov Chain Monte Carlo, Stochastic Simulation for Bayesian Inference*, Chapman and Hall.

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

Geman, S. & Geman, D., 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Analysis and Machine Intelligence* **6**, 721-741.

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.

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.

Haario, H., Saksman, E. & Tamminen, J., 1001. An adaptive Metropolis algorithm. *Bernoulli* **7**(2), 223-242.

Hansen, T.M., Cordua, K.S. & Mosegaard, K., 2012. Inverse problems with non-trivial priors: efficient solution through sequential Gibbs sampling. *Computer & Geosciences* **16**, 593-611.

Hansen, T.M., Cordua, K.S., Zuino, A. & Mosegaard, K., 2016. Probabilistic integration of geo-Information, in integrated imaging of the Earth: Theory and Applications (eds M. Moorkamp, P. G. Lelivre, N. Linde and A. Khan) *John Wiley & Sons, Inc*, Hoboken, NJ.

Hastings, W.K., 1970. Monte Carlo sampling method using Markov chains and their applications. *Biometrika* **57**, 97-109.

Heimsath, A.M., Dietrich, W.E., Nishiizumi, K. & Finkel, R.C., 1997. The soil production function and landscape equilibrium. *Nature* **388**, 358-361.

Holbrook, W.S., Bacon, A.R., Brantley, S.L., Carr, B.J., Flinchum, B.A., Marcon, V., Richter, D. deB. & Riebe, C.S., 2017. Links between physical and chemical weathering inferred from a 65-m-deep borehole through Earth's critical zone. (*Holbrook, personal communication*), July 2017.

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

Iglesias, M.A., Lin, K. & Stuart, A.M., 2014. Well-posed Bayesian geometric inverse problems arising in

subsurface flow. *Inverse Problems* **10**, 114001-114040.

Jongmans, D. & Garambois, S., 2007. Geophysical investigation of landslides: a review. *Bulletin Société Gologique de France* **178**(2), 101-112.

Juhojuntti, N. & Kamm, J., 2015. Joint inversion of seismic refraction and resistivity data using layered models Applications to groundwater investigation *Geophysics* **80**(1), 43-55.

Kabanikhin, S.I. 2008. Definitions and examples of inverse and ill-posed problems. *Journal of Inverse and Ill-Posed Problems* **16**(4), 317-357.

Kim, K.-Y., Seong, H., Kim, T., Park, K.-H., Woo, N.-C., Park, Y.-S., Koh, G.-W. & Park, W.-B., 2006. Tidal effects on variations of fresh-saltwater interface and groundwater flow in a multilayered coastal aquifer on volcanic island (Jeju Usland, Korea). *Journal of Hydrology* **330**, 525-542.

Kneisel, C., Hauck, C. & Moorman, B., 2008. Advances in geophysical methods for permafrost investigations. *Permafrost and Periglacial Processes* **19**, 157-178.

Koefoed O., 1979. *Geosounding Principles I, Resistivity Sounding Measurements.* Elsevier Science Publishing Co., Amsterdam.

Laloy, E., Linde, N., Jacques, D. & Mariethoz, G., 2016. Merging parallel tempering with sequential geostatistical resampling for improved posterior exploration of high-dimensional subsurface categorical fields. *Advances in Water Resources* **90**, 57-69.

Laloy, E. & Vrugt, J.A., 2012. Highdimensional posterior exploration of hydrologic models using multipletry DREAM(ZS) and highperformance computing. *Water Resources Research* **48**, W01526.

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.

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

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. & Teller, E., 1953. Equations of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1097-1092.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory.* International Geophysics Series Volume 45, Academic Press Inc.

Minsley, B.J., Abraham, J.D., Smith, B.D., Cannia, J.C., Voss, C.I., Jorgenson, M.T., Walvoord, M.A., Wylie, B.K., Anderson, L., Ball, L.B., Deszcz-Pan, M., Wellman, T.P. & Ager, T.A., 2012. Airborne electromagnetic imaging of discontinuous permafrost. *Geophysical Research Letters* **39**(2), 1-8.

Moon, S., J. T. Perron, S. J. Martel, W. S. Holbrook, and & St Clair, J. 2017. A model of three-dimensional topographic stresses with implications for bedrock fractures, surface processes, and landscape evolution. *Journal of Geophysical Research - Earth Surface* **122**(4), 823-846.

Mosegaard, K & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. *J. Geophys.*

*Res.* **B7**, 12431-12447.

Nguyen, F., Kemna, A., Antonsson, A., Engesgaard, P., Kuras, O., Ogilvy, R., Gisbert, J., Jorreto, S. & Pulido-Bosch, A., 2009. Characterization of seawater intrusion using 2D electrical imaging. *Near Surface Geophysics*, **7**(5-6), 377-390.

Oldenburg, F.D. & Li, Y., 1999. Estimating depth of investigation in dc resistivity and IP surveys. *Geophysics* **64**(2), 403-416.

Parker, R.L., 1984. The inverse problem of resistivity sounding. *Geophysics* **49** 2143-2158.

Parsekian, A.D., Singha, K., Minsley, B.J., Holbrook, W.S., & Slater, L., 2015. Multiscale geophysical imaging of the critical zone. *Reviews of Geophysics* **53**(1), 1-26.

Rempe, M.D. & Dietrich, W.E., 2014. A bottom-up control on fresh-bedrock topography under landscapes. *PNAS* **111**(18), 6576-6581.

Reneau, S.L. & Dietrich, W.E., 1987. Size and location of colluvial landslides in a steep forested landscape. *Proceedings of the Corvallis Symposium* IAHS Publ.**165**, 39-48.

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., 2011. Advanced Electrical Resistivity Modelling and Inversion using Unstructured Discretization. *PhD thesis* University of Leipzig.

Rücker, C., Günther, T. & Spitzer, K., 2006. Three-dimensional modelling and inversion of dc resistivity data incorporating topography-I. Modelling. *Geophysical Journal International* **166**, 495-505.

Saas, O., 2007. Bedrock detection and talus thickness assessment in the European Alps using geophysical methods. *Journal of Applied Geophysics* **3**(62), 254-269.

Sambridge, M., Gallagher, K., Jackson, A. & Rickwood, P., 2006. Trans-dimensional inverse problems, model comparison and the evidence. *Geophysical Journal International* **167**(2), 528-542.

Sambridge, M. & Mosegaard, K., 2002. Monte Carlo methods in geophysical inverse problems. *Rev. Geophys.* **40**, 3.1-3.29.

Sanchez, G., Rolland, Y., Corsini, M., Braucher, R., Bourls, D., Arnold, M. & Aumaître, G., 2010. Relationships between tectonics, slope instability and climate change: Cosmic ray exposure dating of active faults, landslides and glacial surfaces in the SW Alps. *Geomorphology* **117**(1-2), 1-13.

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 Model Parameter Estimation*. Society for Industrial and

Applied Mathematics.

Tarantola, A. & Valette, B., 1982. Generalized non-linear problems solved using the least-squares criterion. *Review of Geophysics* **20**, 219-232.

Tarantola, A. & Valette, B., 1982. Inverse problem==quest for information. *Geophysics* **50**, 150-170.

Tavassoli, Z., Carter, J. N. & King, P. R., 2015. An analysis of history matching errors. *Computational Geosciences* **9**, 99-123.

Tikhonov, A.N., 1963. Regularization of ill-posed problems. *Doklady Akademii Nauk SSSR* **153**, 1-6.

Uhlemann, S., Chambers, J., Wilkinson, P., Maurer, H., Merritt, A., Meldrum, P., Kuras, O., Gunn, D., Smith, A. & Dijkstra, T., 2017. Four-dimensional imaging of moisture dynamics during landslide reactivation. *Journal of Geophysical Research: Earth Surface* **122**, 398-418.

Vignoli, G., Sapia, V., Menghini, A., & Viezzoli, A., 2017. Examples of improved inversion of different airborne electromagnetic datasets via sharp regularization. *Journal of Environmental and Engineering Geophysics* **22**(1), 51-56.

Ward, W.O.C., Wilkinson, P.B., Chambers, J.E., Oxby, L.S. & Bai, L., 2014. Distribution-based fuzzy clustering of electrical resistivity tomography images for interface detection. *Geophysical Journal International* **197**, 310-321.

Werner, A.D., Bakker, M., Post, V.E.A., Vandenbohede, A., Lu, C., Ataie-Ashtiani, B., Simmons, C.T. & Barry, D.A., 2013. Seawater intrusion processes, investigation and management: Recent advances and future challenges. *Advances in Water Resources* **51**, 3-26.

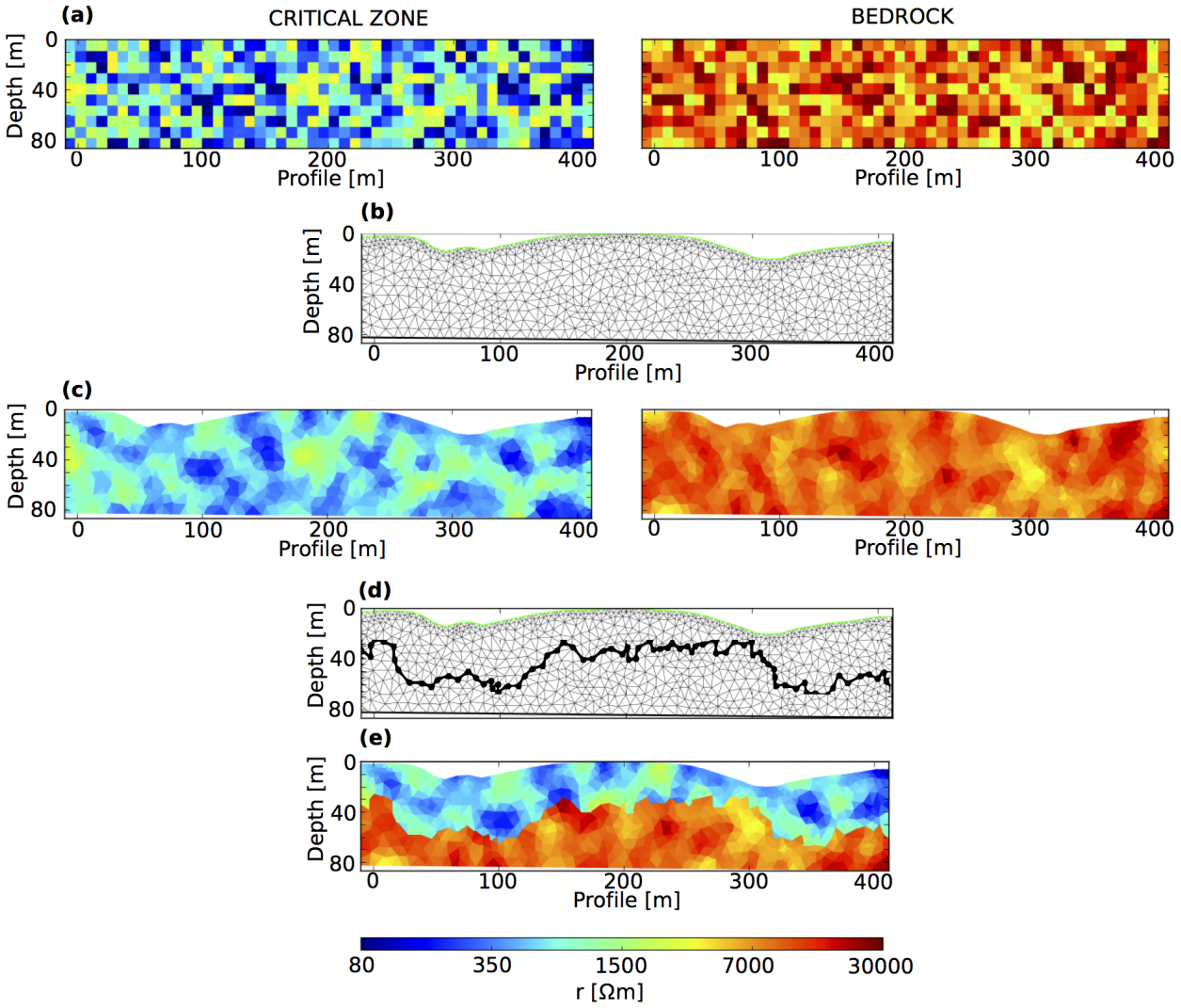


**Figure 1.** Proposed model setting for inversion of geophysical datasets in the presence of an unknown interface, which divides the investigated subsurface in two sub-domains. (a) Physical properties, representative of the CZ ($\mathbf{m}_{CZ}$) and bedrock ($\mathbf{m}_b$) in the present example, defined on regular grids over the whole domain of investigation. (b) Unstructured mesh used for forward computation and (c) the corresponding interpolation of the regular grids (a) onto this mesh. (d) The interface $\mathbf{I}$, defined by connected nodes of the forward computation mesh, is used to assign (e) the electrical properties used in the forward calculations for a given model ($\mathbf{m} = G(\mathbf{m}_{CZ}, \mathbf{m}_b, \mathbf{I})$).

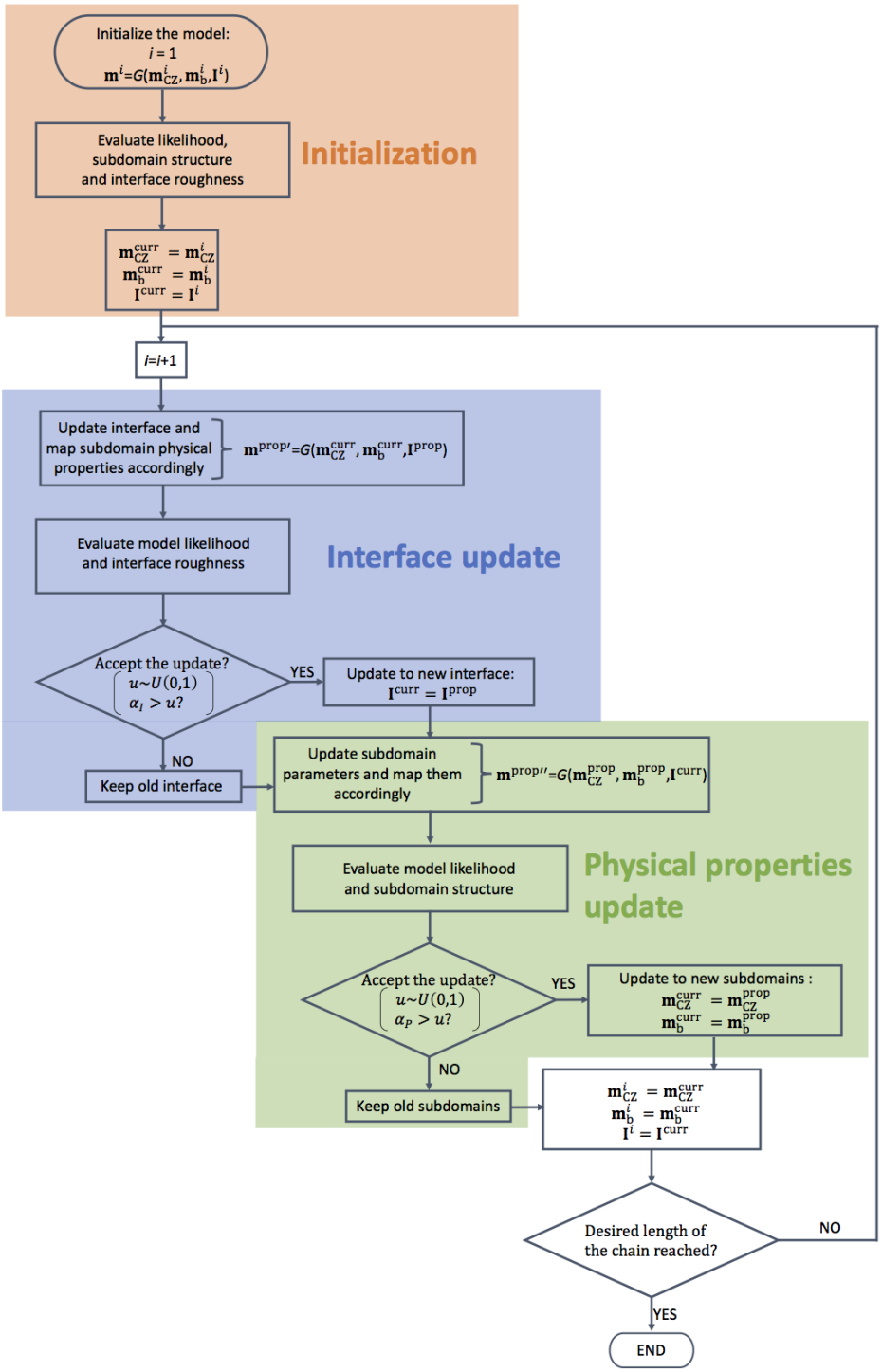


**Figure 2.** Flow chart of the proposed empirical-Bayes-within-Gibbs algorithm. First, we initialize the model (as described in Fig. 1) and evaluate its likelihood, physical properties structure and interface roughness. We then sequentially evaluate proposed interface and physical property updates in an iterative scheme until the chosen length of the chain is reached. Here, $u \sim U(0,1)$ represents a randomly chosen number with uniform probability between 0 and 1, while the acceptance ratios $\alpha_I$ and $\alpha_P$ are described in eqs. 7 and 8, respectively.

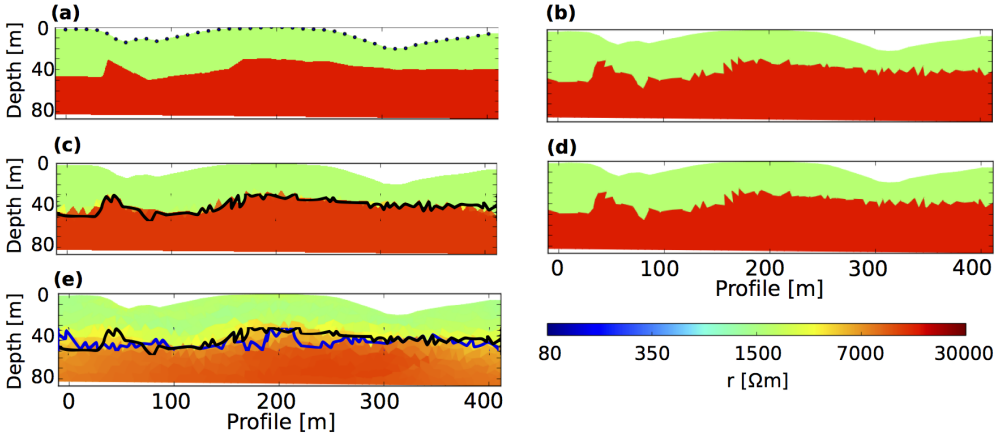


**Figure 3.** (a) Resistivity model and electrode positions (black dots) used to generate the synthetic dataset with homogeneous sub-domains and (b) interpolation of (a) on the unstructured mesh used for ERT forward modeling. (c) Mean of posterior realizations from the empirical-Bayes-within-Gibbs chain and (d) model realization corresponding to the maximum likelihood. (e) Deterministic inversion result. In (c) and (e), the black line indicates the interface we aim to infer, while the blue line in (e) is the interface obtained using the maximum vertical gradient in the deterministic inversion result.

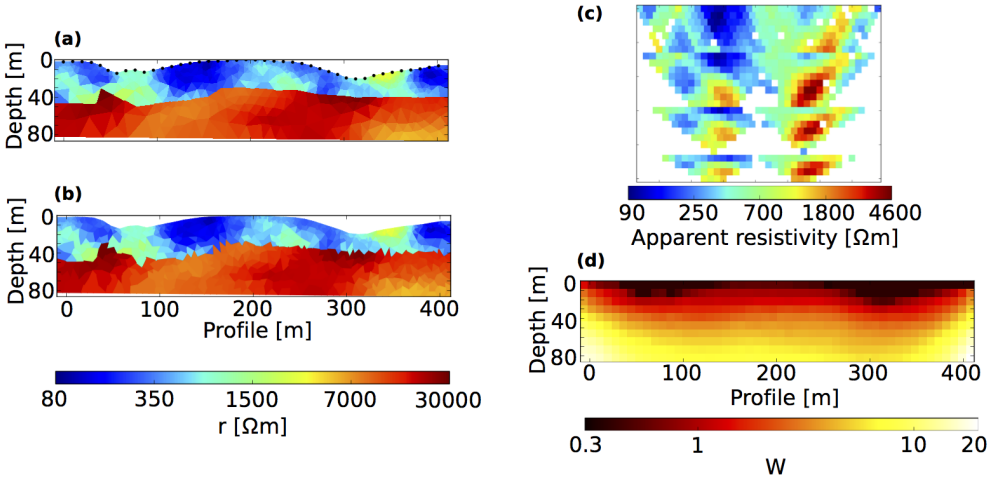


**Figure 4.** (a) Resistivity model and electrode positions (black dots) used to generate the synthetic dataset with heterogeneous sub-domains and (b) interpolation of (a) on the unstructured mesh used for ERT forward modeling. (c) Noise-contaminated synthetic dataset obtained from (a) (645 apparent resistivity data points). (d) Distributed weights applied to the model parameter update.

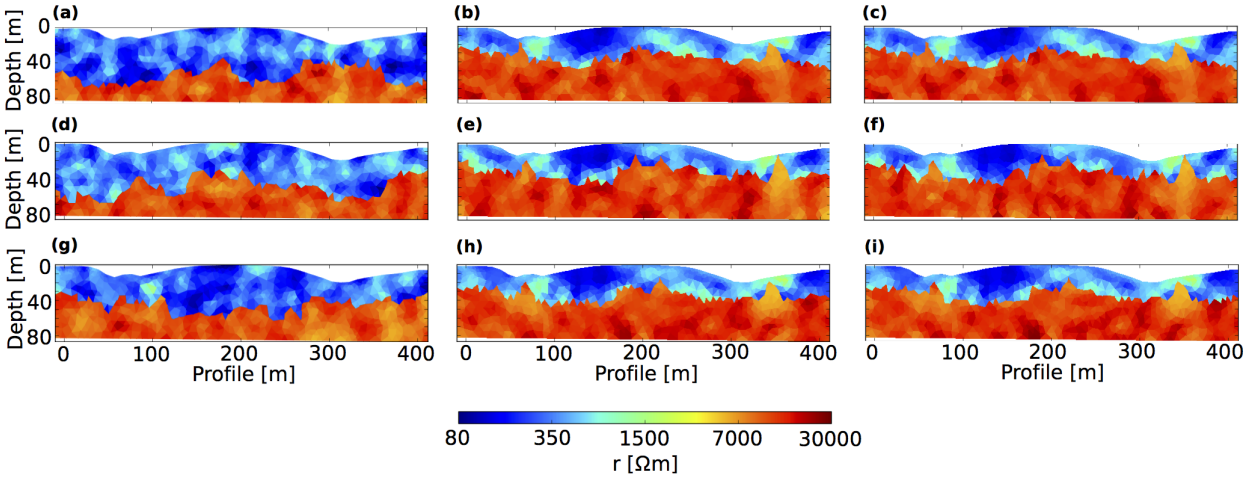


**Figure 5.** (a), (d) and (g) The three initial models used for the synthetic test case with heterogeneous sub-domains. (b), (e) and (h) The last corresponding realizations of each MCMC chain. (c), (f) and (i) Corresponding maximum likelihood realizations.

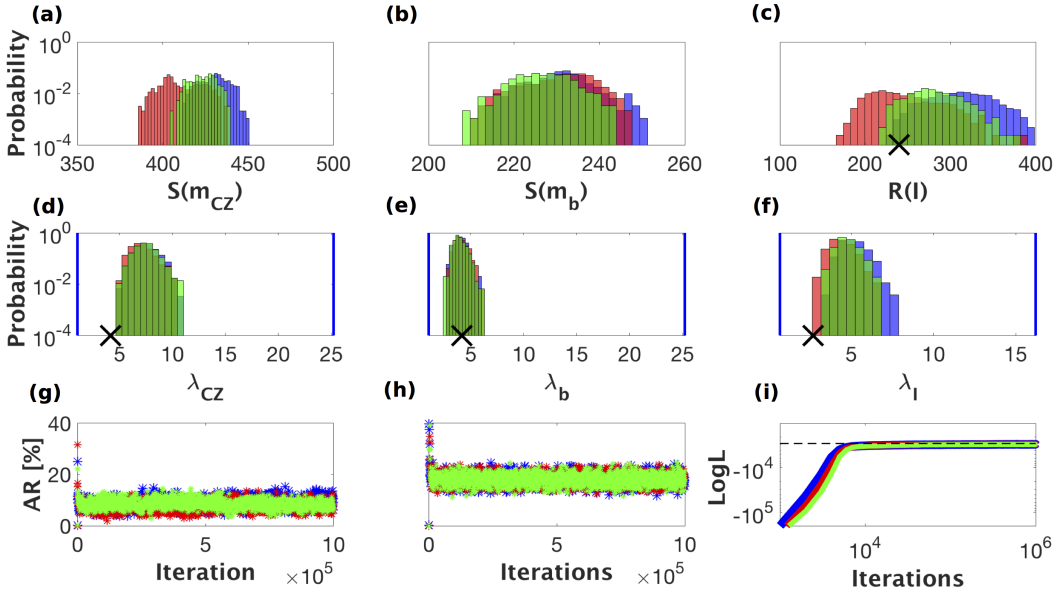


**Figure 6.** Behavior of the three independent chains (each colors refers to a specific chain) for the synthetic example with heterogeneous sub-domains. Sampled model structure for (a) CZ, (b) bedrock and (c) regolith-bedrock interface, with the black cross referring to the values for the actual interface. Sampled hyper parameters penalizing (d) CZ, (e) bedrock and (f) interface roughness, with black crosses referring to the optimal hyper parameter values for the true model. Acceptance rate of (g) the interface and (h) the physical property update. (a)-(f) share the same y-label, so does (g) and (h). (i) Log likelihood evolution during the course of the inversion with target log-likelihood (dashed black line)

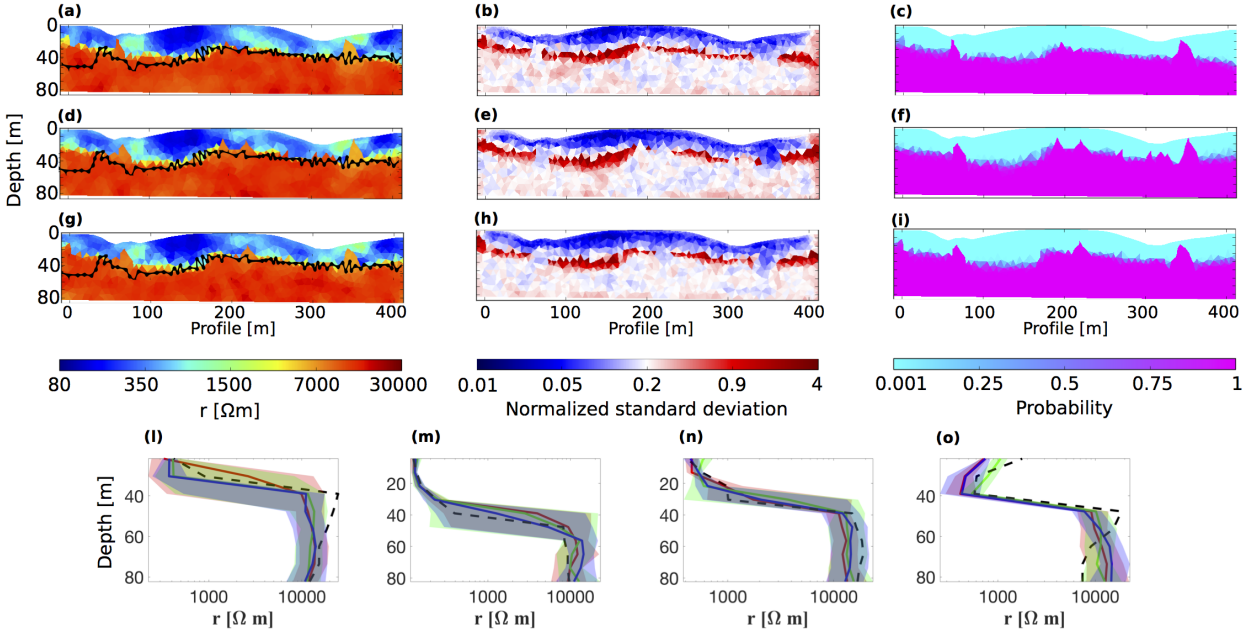


**Figure 7.** Posterior model results for empirical-Bayes-within-Gibbs inversion of the synthetic dataset with heterogeneous sub-domains. (a), (d) and (g) represent the mean of the posterior model realizations for each chain with the actual interface superimposed (black line). (b), (e) and (h) show the standard deviation of the electrical resistivity divided by the corresponding mean values. (c), (f) and (i) show the probability of each cell of the mesh being part of the bedrock. (l), (m), (n) and (o) are vertical resistivity profiles at 50, 150, 250 and 350 meters along the profile; we show in blue, red and green the mean model results (solid lines) of each chain with the inferred posterior ranges (shadow areas), while the black dashed lines represent the synthetic model.

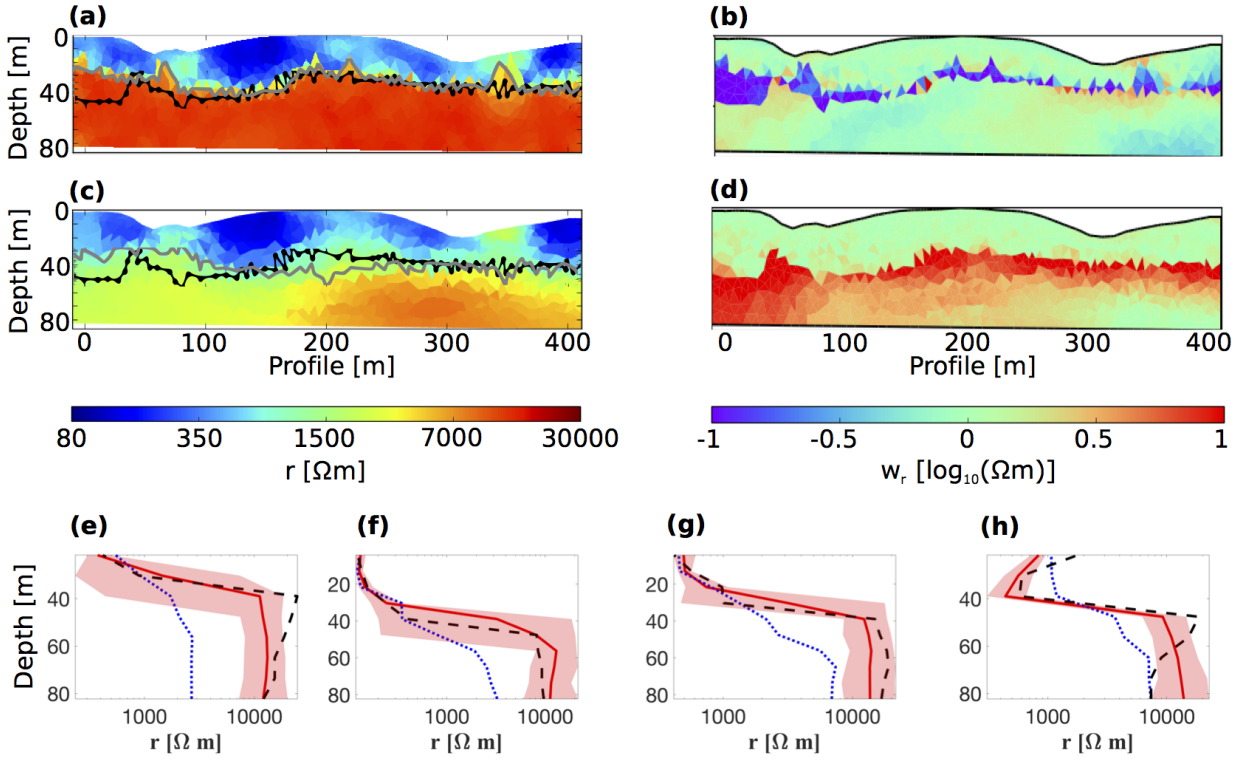


**Figure 8.** (a) Mean of the posterior model realizations for the three empirical-Bayes-within-Gibbs chains and (b) its model discrepancy compared to the true model. (c) Deterministic inversion result and its (d) model discrepancy. The black line in (a) and (c) it indicates the interface we aim to infer (i.e., Fig. 4b), the gray line in (a) is the mean of the inferred posterior interfaces, while in (c) represents the interface obtained with the maximum vertical gradient method based on the deterministic inversion result. (e), (f), (g) and (h) show vertical resistivity profiles respectively at 50, 150, 250 and 350 meters along the profile. The red solid lines represent the mean model of empirical-Bayes-within-Gibbs inversion and the red shadow areas indicate the inferred posterior ranges, the blue dotted lines are obtained from the deterministic inversion, while the black dashed lines represent the underlying synthetic model.

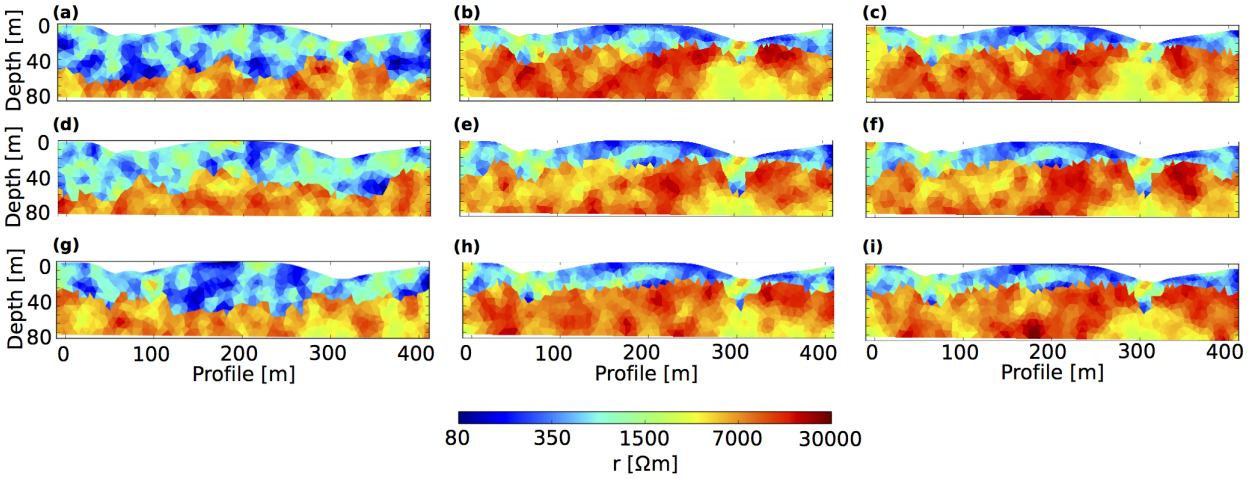


**Figure 9.** (a), (d) and (g) The three initial models used for the field-data inversion using empirical-Bayes-within-Gibbs sampling. (b), (e) and (h) The last corresponding realizations of each MCMC chain. (c), (f) and (i) Corresponding maximum likelihood realizations.

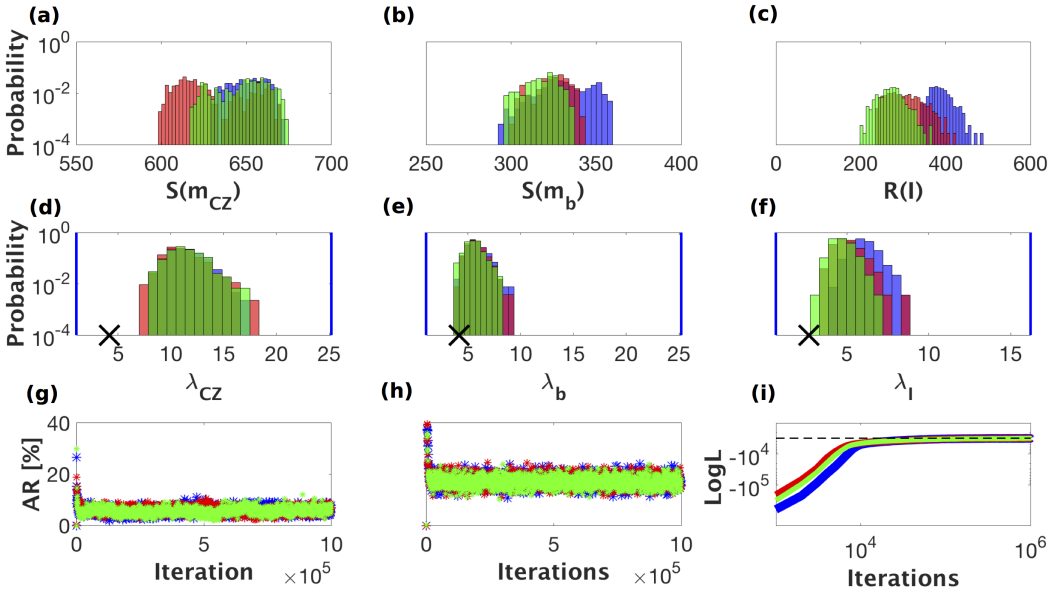


**Figure 10.** Characteristics of the three independent chains (each colors refers to a specific chain) for the field example. Sampled model structure for (a) CZ, (b) bedrock and (c) regolith-bedrock interface. Sampled hyper parameters penalizing (d) CZ, (e) bedrock and (f) interface roughness, with black crosses referring to the optimal hyper parameter values for the true model. Acceptance rate of (g) the interface and (h) the physical property update. (a)-(f) share the same y-label, so does (g) and (h). (i) Log likelihood evolution during the course of the inversion with target log-likelihood (dashed black line)

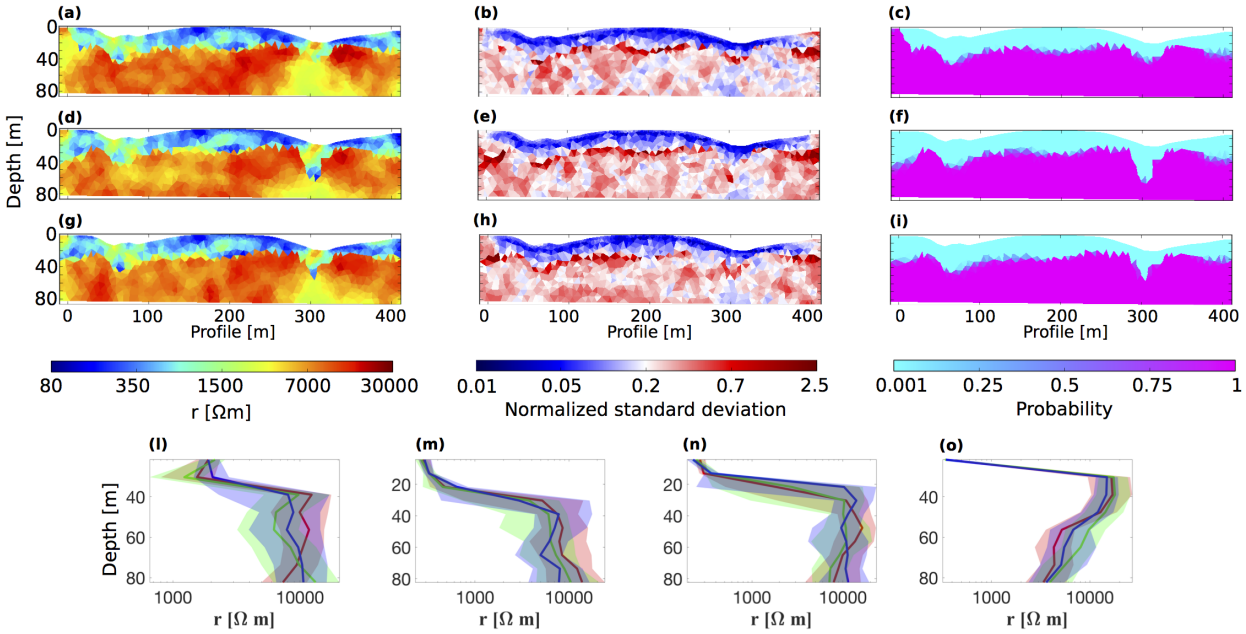


**Figure 11.** Posterior model results for empirical-Bayes-within-Gibbs inversion of the field dataset. (a), (d) and (g) represent the mean of the posterior model realizations for each chain. (b), (e) and (h) show the standard deviation of the electrical resistivities divided by the corresponding mean values. (c), (f) and (i) show the probability for each cell of the forward mesh to be part of the bedrock. (l), (m), (n) and (o) are vertical resistivity profiles at 50, 150, 250 and 350 meters along the profile. Blue, red and green indicate the mean model of each chain (solid lines) with the inferred posterior ranges (shadow areas).

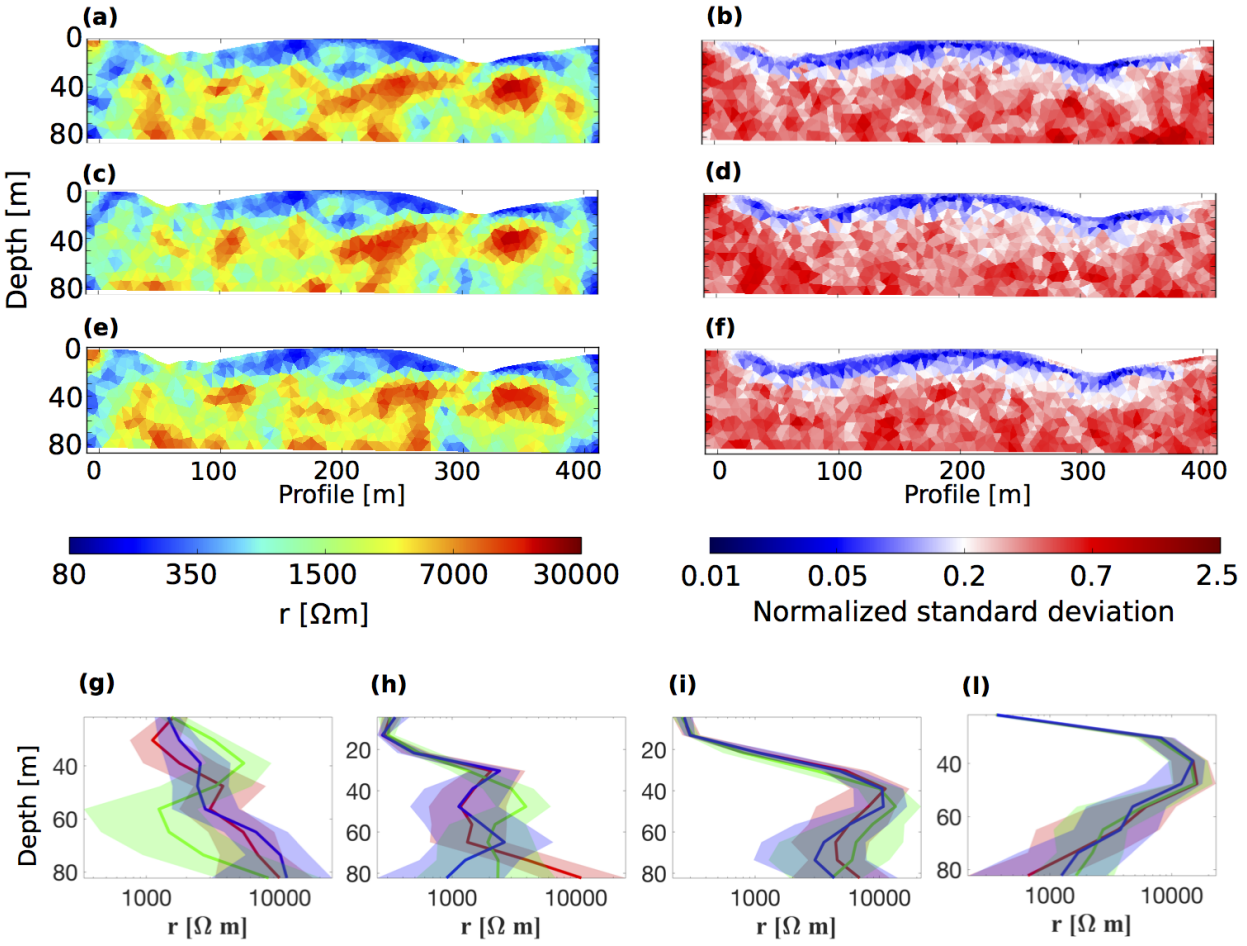


**Figure 12.** Posterior model results for one-domain empirical Bayes inversion of the field dataset. (a), (c) and (e) represent the mean of the posterior model realizations for each chain, while (b), (d) and (f) show the standard deviation of the electrical resistivities divided by the corresponding mean values.(g), (h), (i) and (l), vertical resistivity profiles respectively at 50, 150, 250 and 350 meters along the profile. Blue, red and green indicate the mean model of each chain (solid lines) with the inferred posterior ranges (shadow areas).

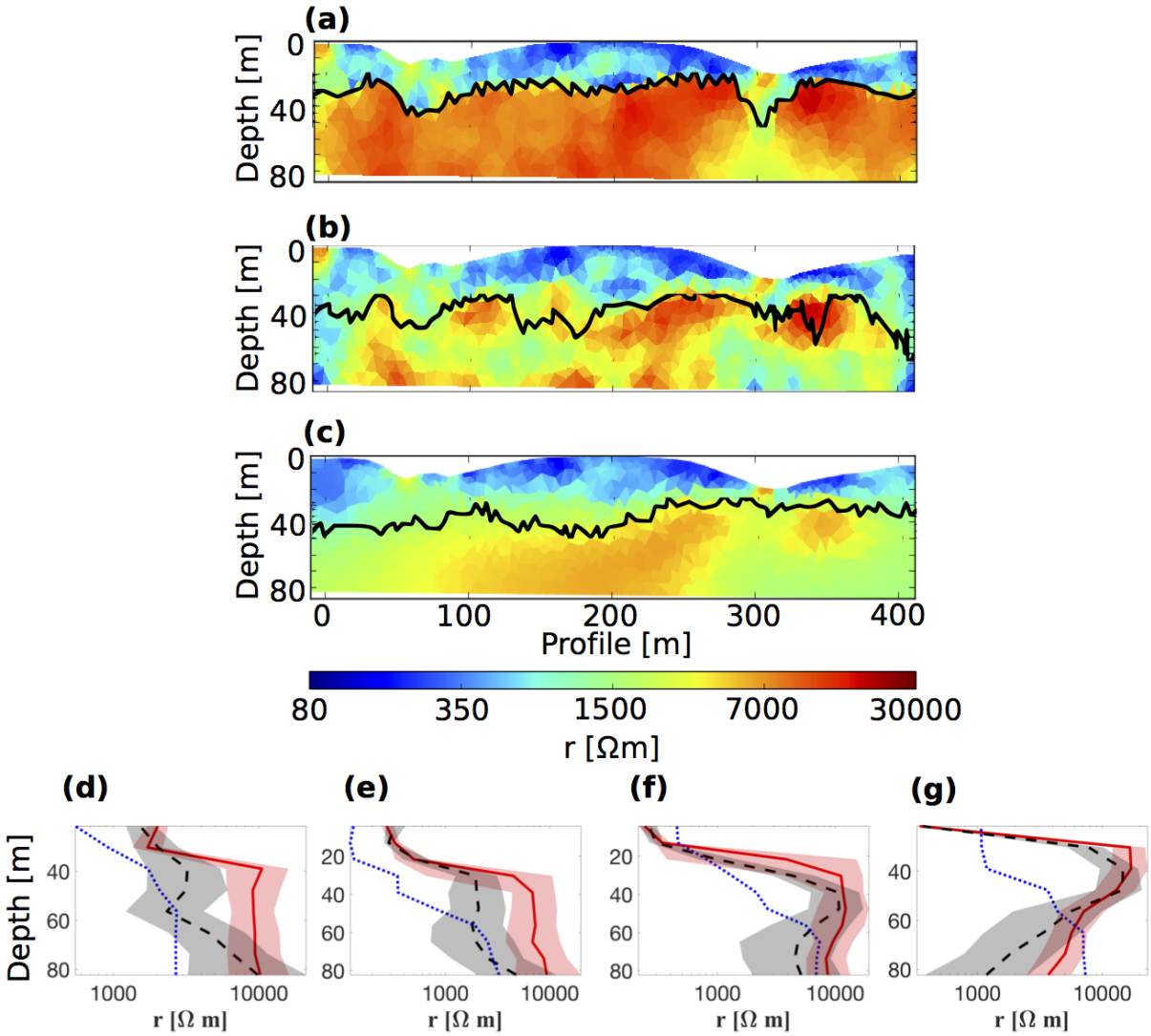


**Figure 13.** Comparison of field data inversion results. (a) Mean of the posterior model realizations for the three empirical-Bayes-within-Gibbs chains, (b) mean of posterior model realizations for the three one-domain empirical Bayes chains and (c) deterministic inversion result. In (a) the black line is the mean of the inferred posterior interfaces, while in (b) and (c) it represents the interfaces obtained by the maximum vertical gradient method. (d), (e), (f) and (g) show vertical resistivity profiles at 50, 150, 250 and 350 meters along the profile. Red solid lines are used to represent the mean of the empirical-Bayes-within-Gibbs inversion result and the red shadow areas the inferred posterior ranges, the black dashed lines with the gray shadow areas represent the one-domain empirical Bayes inversion results and the blue dotted lines are obtained from the deterministic inversion.