A Geostatistical Markov Chain Monte Carlo inversion algorithm for Electrical Resistivity Tomography

Mattia Aleardi°, Alessandro Vinciguerra*, Azadeh Hojat[+,§]

[°]University of Pisa, Earth Sciences Department, via S. Maria 53, 56126, Pisa, Italy

[*]University of Florence, Earth Sciences Department, via. G. La Pira 4, 50121, Florence, Italy

[+]Shahid Bahonar University of Kerman, Department of Mining Engineering, Kerman, 76188, Iran

[§]Politecnico di Milano, Department of Civil and Environmental Engineering, Piazza Leonardo da Vinci 32, 20133, Milano Italy

Corresponding author: Mattia Aleardi, mattia.aleardi@unipi.it

ABSTRACT

Electrical resistivity tomography (ERT) is an ill-posed and non-linear inverse problem commonly solved through deterministic gradient-based methods. These methods guarantee a fast convergence toward the final solution, but the local linearization of the inverse operator impedes accurate uncertainty assessments. On the contrary, numerical Markov Chain Monte Carlo (MCMC) algorithms allow for accurate uncertainty appraisals but appropriate MCMC recipes are needed to reduce the computational effort and make these approaches suitable to be applied to field data. A key aspect of any probabilistic inversion is the definition of an appropriate prior distribution of the model parameters that can also incorporate spatial constraints to mitigate the ill-conditioning of the inverse problem. Usually, Gaussian priors oversimplify the actual distribution of the model parameters that often exhibit multimodality due to the presence of multiple litho-fluid facies. In this work, we develop a novel probabilistic MCMC approach for inversion of ERT data. This approach jointly estimates resistivity values, litho-fluid facies, along with the associated uncertainties from the measured

apparent resistivity pseudosection. In our approach, the unknown parameters include the facies model as well as the continuous resistivity values. At each spatial location, the distribution of the resistivity value is assumed to be multimodal and non-parametric with as many modes as the number of facies. An advanced MCMC algorithm (the Differential Evolution Markov Chain) is used to efficiently sample the posterior density in a high-dimensional parameter space. A Gaussian variogram model and a first-order Markov chain respectively account for the lateral continuity of the continuous and discrete model properties, thereby reducing the null-space of solutions. The Direct Sequential Simulation geostatistical method allows the generation of sampled models that honor both the assumed marginal prior and spatial constraints. During the MCMC walk, we iteratively sample the facies, by moving from one mode to another one, and the resistivity values, by sampling within the same mode. The proposed method is first validated by inverting the data calculated from synthetic models. Then, it is applied to field data and benchmarked against a standard local inversion algorithm. Our experiments prove that the proposed MCMC inversion retrieves reliable estimations and accurate uncertainty quantifications with a reasonable computational effort.

INTRODUCTION

Electrical resistivity is an important property of geological formations with high sensitivity to fluid saturation and porosity and thus, the Electrical Resistivity Tomography (ERT) method is widely and successfully employed for groundwater exploration, geotechnical characterization, mapping of contaminant plumes, and monitoring of landfills and levees (Legaz et al., 2009; Johnson et al., 2010; Müller et al., 2010; Pollock and Cirpka, 2012; Moradipour et al., 2016; Arosio et al., 2017; Whiteley et al., 2017; Crawford et al., 2018; Fiandaca et al., 2018; Hojat et al., 2019a; Tresoldi et al., 2019; Siemon et al., 2019; Hermans and Paepen, 2020). The ERT inverse problem is nonlinear and ill-posed and is usually solved through deterministic gradient-based algorithms (Zhang et al., 2005; Pidlisecky and Knight, 2008; Karoulis et al., 2014) that linearize the problem around an initial solution thereby losing the information for accurate uncertainty appraisals. Moreover, these methods are prone to get

trapped in local minima of the misfit function and thus, the starting model must lie within the global minimum valley. Markov Chain Monte Carlo algorithms overcome these limitations and can be employed to cast nonlinear inverse problems into a solid Bayesian framework in which the final solution is the so-called posterior probability density (PPD) function in model space (Kaipio, 2000; Sambridge and Mosegaard, 2002; Tarantola, 2005; Ramirez et al., 2005; Linde et al., 2006; Grana and Della Rossa, 2010; Irving and Singha, 2010; Sen and Stoffa, 2013; Jardani et al., 2013; Aleardi et al., 2017; Aleardi and Mazzotti, 2017; Galetti and Curtis, 2018; Andersen et al., 2018; Aleardi and Salusti, 2020a) that fully quantifies the ambiguities in the recovered model. Although the increased computational power provided by modern parallel architectures has considerably encouraged the applications of MCMC methods to solve geophysical problems, it is always crucial adopting specific recipes to guarantee an accurate and computationally efficient sampling of the PPD. For example, many MCMC algorithms (e.g. the popular random walk Metropolis) are known to mix slowly between the modes if the target distribution is multimodal with modes separated by low probability regions (Holmes et al., 2017; Aleardi et al., 2018; Scalzo et al., 2019). A simple approach to mitigate this issue makes use of multiple MCMC chains to sample the PPD. This strategy usually offers robust protection against premature convergence because the chains use different trajectories to explore the parameter space. However, this approach is inefficient in high-dimensional problems where the curse-of-dimensionality makes the target distribution highly localized within each model space dimension. This issue usually increases the probability for the MCMC chains to get trapped in local maxima of the PPD. Therefore, a considerable number of samples is needed to achieve accurate posterior estimations. There have been many attempts to improve the convergence of MCMC algorithms in case of high-dimensional problems, for example hybridizing standard MCMC algorithms with global search methods (e.g. Differential evolution Markov Chain “DEMC”, or Differential evolution adaptive Metropolis; Turner and Sederberg, 2012; Vrugt, 2016).

One challenge often posed by geophysical inversions concerns the simultaneous estimation of discrete (i.e. litho-fluid facies) and continuous (e.g. resistivity values) model parameters from the

observed data (Gunning and Sams, 2018). Another challenge is related to the complexity of the property distribution and spatial correlation. On one hand, spatial model constraints can be used to reduce the ill-conditioning of the inversion problem (Doyen, 2007) and geostatistical simulation algorithms are routinely employed for this purpose (Azevedo and Soares, 2017). For example, the well-known simple kriging approach can be used for simple Gaussian priors, while in case of non-Gaussian assumptions more sophisticated approaches (e.g. the Direct Sequential Simulation "DSS", Soares, 2001; Horta and Soares, 2010) are needed. On the other hand, the actual distribution of the resistivity values in the subsurface is often multimodal due to the presence of multiple litho-facies (e.g. clay and sand). In these contexts, a Gaussian or even Gaussian-mixture prior model could not be adequate to reliably capture the complex relations among the resistivity values and the litho-facies, and for this reason, non-parametric priors should be employed. In this context, one of the outstanding benefits of MCMC algorithms is the possibility to reliably assess the PPD in case of complex mixed discrete-continuous problems with non-Gaussian priors.

In this work, we implement a 2D MCMC approach for ERT inversion that simultaneously estimates resistivity values, litho-fluid facies, together with the related uncertainties from the apparent resistivity pseudosection. This approach solves this mixed continuous-discrete inversion problem by adopting a 2.5D finite elements forward operator (Karoulis et al., 2013) and assuming a non-parametric mixture prior, in which each mode is associated with a specific facies. A 2D Gaussian variogram model defines the spatial correlation pattern of the resistivity values, whereas a first-order Markov chain preserves the lateral and vertical continuity of the discrete property distribution (Ulvmoen and Omre, 2010). Multiple facies (e.g., different geological formations characterized by different mineralogical compositions and/or water content) can be found in ERT application. In the context of MCMC algorithms, the simultaneous presence of continuous and discrete parameters may result in persistent rejections of models, a very low acceptance ratio, and slow convergence of the sampling toward stable PPD estimations. Therefore, we adopt a specific MCMC recipe specifically designed for multimodal distributions and mixed discrete-continuous problems (also see Holmes et

al., 2017; and Aleardi and Salusti, 2020a). Our recipe includes a facies move and a resistivity move. In the former, the current facies model is perturbed thus jumping from one mode to another one, while in the latter, only the continuous property values are perturbed thus sampling within the same mode. The Differential Evolution Markov Chain (DEMC) method is employed to efficiently sample the PPD in a high-dimensional parameter space, whereas the DSS algorithm is used to generate sampled models that honor both the prior assumption and the assumed spatial correlation pattern. The methodology is first assessed on synthetic data and then applied to field data.

THE IMPLEMENTED MCMC ALGORITHM

MCMC algorithms generate subsurface models that honor the prior information and accept or reject these models based on their posterior likelihood value. If $\mathbf{m}_{\boldsymbol{curr}}$ is the current model and $\mathbf{m}_{\boldsymbol{prop}}$ is the proposed (perturbed) model, the probability for the chain to move from $\mathbf{m}_{\boldsymbol{curr}}$ to $\mathbf{m}_{\boldsymbol{prop}}$ is defined by the Metropolis-Hasting rule:

$$p(\mathbf{m}_{prop}|\mathbf{m}_{curr}) = \min\left[1, \frac{p(\mathbf{m}_{prop})}{p(\mathbf{m}_{curr})} \times \frac{p(\mathbf{d}|\mathbf{m}_{prop})}{p(\mathbf{d}|\mathbf{m}_{curr})} \times \frac{q(\mathbf{m}_{curr}|\mathbf{m}_{prop})}{q(\mathbf{m}_{prop}|\mathbf{m}_{curr})}\right], \quad (1)$$

where $\mathbf{d}$ is the observed data vector, $q()$ is the proposal distribution that defines the new model $\mathbf{m}_{prop}$ as a random deviate from a probability distribution $q(\mathbf{m}_{prop}|\mathbf{m}_{curr})$ conditioned only on the current model $\mathbf{m}$. Note that the proposal ratio term in equation (1) vanishes if symmetric proposals (for example a Gaussian proposal) are employed. If $\mathbf{m}_{prop}$ is accepted $\mathbf{m}_{curr} = \mathbf{m}_{prop}$ and another model is generated as a random perturbation of $\mathbf{m}_{curr}$. Otherwise $\mathbf{m}_{curr}$ is retained and another proposed model is generated. The ensemble of accepted models after the burn-in period is used to numerically compute the PPD.

The DEMC is an advanced MCMC algorithm that uses a population of different chains that evolve according to differential evolution principles (Ter Braak, 2006). In more detail, such differential evolution principles are used to generate multivariate proposals for each DEMC chain: let the *m*-vector $\mathbf{m}$ represent the state of a single chain, then at each iteration *t*-1, the Q chains define a

population $\mathbf{M} = \{\mathbf{m}_{t-1}^1, \mathbf{m}_{t-1}^2, \ldots, \mathbf{m}_{t-1}^Q\}$ which corresponds to a $Q \times m$ matrix. Multivariate proposals $\mathbf{m}_p$ are defined as:

$$\mathbf{m}_p^i = \mathbf{m}_{t-1}^i + \gamma\left(\mathbf{m}_{t-1}^a - \mathbf{m}_{t-1}^b\right) + \epsilon, \quad a \neq b \neq i \quad (2)$$

where i is the index of the current chain, γ denotes the jump rate, a and b are integer values drawn without replacement from $\{1, \ldots, i\text{-}1, i\text{+}1, \ldots, Q\}$, and ϵ represents a small random perturbation drawn from a normal distribution with a small standard deviation σ tailored to the problem at hand: $\epsilon = \mathcal{N}(0, \sigma)$. Each proposal is accepted with Metropolis probability (see equation 1). If the proposal is accepted $\mathbf{m}_t^i = \mathbf{m}_p^i$, otherwise $\mathbf{m}_t^i = \mathbf{m}_{t-1}^i$. The optimal γ parameter depends on the model dimensionality and on the shape of the posterior. For example, for Gaussian assumptions it is usually set to $\gamma = 2.38/2d$, where d represents the number of model parameters. Besides, with a 10% probability, the value of $\gamma = 1$ allows mode-jumping which is a significant strength of DEMC compared to more conventional MCMC methods (i.e. random walk Metropolis or adaptive Metropolis). Different from standard random walk Metropolis, the DEMC also avoids selection of an appropriate proposal distribution because it automatically adapts the proposal to the spread of the PPD associated with each model parameter. This spread is directly inferred from the difference $\mathbf{m}_{t-1}^a - \mathbf{m}_{t-1}^b$ (see equation 2). Additional and more detailed theoretical insights into the DEMC, together with a Matlab implementation can be found in Vrugt (2016).

Determining the convergence of the sampling to a stable posterior model is a crucial aspect of any MCMC inversion and to this end, we analyze the evolution of the potential scale reduction factor (PSRF) for the continuous model parameters (Gelman et al., 1995). We remind that a PSRF lower than 1.2 for a given unknown proves that convergence has been achieved for that model parameter.

Let us now focus on the implemented MCMC-ERT inversion in which $\boldsymbol{\pi}$ and $\mathbf{e}$ represent the facies and the resistivity values, respectively. In our case of a mixed discrete-continuous inverse problem the posterior model can be derived from the chain rule:

$$p(\mathbf{e}, \boldsymbol{\pi}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{e}, \boldsymbol{\pi})p(\mathbf{e}|\boldsymbol{\pi})p(\boldsymbol{\pi})}{p(\mathbf{d})} = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})} \quad (3)$$

where $\mathbf{m} = [\mathbf{e}, \boldsymbol{\pi}]$. Before the MCMC inversion, the available borehole data and/or geological information about the investigated area must be exploited to define the $p(\boldsymbol{\pi})$ and $p(\mathbf{e}|\boldsymbol{\pi})$ distributions. The former is the discrete prior distribution of facies, whereas the latter is the prior distribution of the resistivity values within each facies. In our implementation, the $p(\mathbf{e}|\boldsymbol{\pi})$ is a non-parametric mixture model that can be derived, for example, by using the kernel density estimation algorithm. The lateral correlation of the continuous property is modeled according to a Gaussian variogram model that along a single spatial direction (e.g., the horizontal *x*-direction) is defined by the following correlation function:

$$\tau_x = \exp\left(-\frac{h_x^2}{\alpha_x^2}\right), \qquad (4)$$

where h_x is the spatial distance of the autocorrelation function along the *x*-direction, and α_x is the parameter that defines the spatial dependency along the same *x*-axis. Note that the scalar value α_x directly influences the variogram range along the *x*-direction. To promote the lateral continuity of the discrete property, we employ a first-order Markov model simulation during the MCMC sampling in which the probability of a transition from facies *b* to facies *a* at the spatial position with horizontal and vertical coordinates x and y, respectively, is computed as follows:

$$p\left(\boldsymbol{\pi}_{x,y}^{a} \middle| \boldsymbol{\pi}_{x-1,y}^{h}, \boldsymbol{\pi}_{x+1,y}^{h}, \boldsymbol{\pi}_{x,y-1}^{h}, \right) = p\left(\boldsymbol{\pi}_{x-1,y}^{h} \middle| \boldsymbol{\pi}_{x,y}^{a}\right) p\left(\boldsymbol{\pi}_{x+1,y}^{h} \middle| \boldsymbol{\pi}_{x,y}^{a}\right) p\left(\boldsymbol{\pi}_{x,y-1}^{h} \middle| \boldsymbol{\pi}_{x,y}^{a}\right), \quad (5)$$

where the coordinates $(x-1, y)$, $(x+1, y)$, and $(x, y-1)$ correspond to the first three neighboring positions surrounding the considered location (x, y), whereas the superscript *h* indicates the actual facies configuration at each neighboring position. The probability to move from facies *b* at the neighboring positions $(x-1, y)$ and $(x+1, y)$ to facies *a* at position (x, y) can be derived from the *b*-th row and *a*-th column of the lateral transition matrix $\mathbf{T}^l$. For example:

$$p(\boldsymbol{\pi}_{x,y}^{a} | \boldsymbol{\pi}_{x-1,y}^{b}) = \mathbf{T}_{b,a}^{l}. \qquad (6)$$

Similarly, the probability to move from facies *b* at the neighboring positions $(x, y-1)$ to facies *a* at position (x, y) can be derived from the *b*-th row and *a*-th column of the vertical transition matrix $\mathbf{T}^v$:

$$p(\boldsymbol{\pi}_{x,y}^{a} | \boldsymbol{\pi}_{y-1}^{b}) = \mathbf{T}_{b,a}^{v}. \qquad (7)$$

After defining the prior model and the transition matrices, the acquired data must be exploited to sample from the PPD. The following steps are advocated to this task:

1) Starting model generation: for the *Q* DEMC chains use the DSS algorithm to generate current models $\mathbf{m}_{curr}^{Q} = [\mathbf{e}_{curr}^{1}, \mathbf{e}_{curr}^{2}, \dots, \mathbf{e}_{curr}^{Q}, \boldsymbol{\pi}_{curr}^{1}, \boldsymbol{\pi}_{curr}^{2}, \dots, \boldsymbol{\pi}_{curr}^{Q}]$ for the first MCMC iteration that honor the prior distributions $p(\mathbf{e}|\boldsymbol{\pi})$ and $p(\boldsymbol{\pi})$. Afterward, run the forward modeling and compute the data $\mathbf{d}_{pre}$ associated with each generated model.
2) For each chain randomly select a spatial position $i_q(x, y)$ to be perturbed, where the subscript *q* indicates the *q*-th chain.
3) For each chain draw a random number $\boldsymbol{\kappa_q}$ uniformly distributed over [0,1] and set the $\boldsymbol{\xi}$ parameter defined over [0,1]. Then, loop over the chains and perform the following steps:

   3.1) If $\boldsymbol{\kappa_q} \leq \xi$, apply the *facies move*: perturb the current facies model $\boldsymbol{\pi}_{\boldsymbol{curr}}^{\boldsymbol{q}}$ at the selected position $\boldsymbol{i_q}(\mathbf{x}, \boldsymbol{y})$, thus deriving the proposed facies model $\boldsymbol{\pi}_{\boldsymbol{prop}}^{\boldsymbol{q}}$. The probability of replacing the current facies at the selected position depends on the previously defined vertical and lateral transition matrices (see equation 5). Then, the proposed resistivity value at the i_q-th position is a random realization from the prior distribution $\boldsymbol{p}(\mathbf{e}|\boldsymbol{\pi}_{\boldsymbol{prop}}^{\boldsymbol{q}})$.

   3.2) If $\boldsymbol{\kappa_q} > \boldsymbol{\xi}$, apply the *resistivity move*: preserve the current facies at the i_q -th location (so that $\boldsymbol{\pi}_{\boldsymbol{prop}}^{\boldsymbol{q}} = \boldsymbol{\pi}_{\boldsymbol{curr}}^{\boldsymbol{q}}$) and perturb only the resistivity values according to the DEMC strategy (see equation 2) in which the algorithm selects other two chains with indexes *a* and *b* that share with the chain *q* the same facies at the i_q-th position. If there are no chains *a* and *b* that share with the chain *p* the same facies at the i_q -th position, apply a simple random walk Metropolis update in which the perturbation follows a Gaussian proposal with a zero mean and a previously defined covariance matrix $\boldsymbol{\mathcal{N}}(\mathbf{e}; \mathbf{0}; \sum_{\boldsymbol{p}})$.
4) *DSS updating*: Use the previously defined variogram model and the DSS algorithm to propagate the perturbation of the resistivity value at the i_q-th position over the neighboring

spatial locations that are determined according to the ranges of the 2-D variogram model. All the resistivity values at the spatial locations falling beyond the ranges and at the previously perturbed i_q-th position are excluded from this neighboring ensemble and considered as hard data in the geostatistical simulation process. It results that only the resistivity values at the neighboring positions are perturbed (see Figure 1). This step is crucial to ensure that all the sampled resistivity models honor both the assumed spatial variability pattern and the marginal priors. At the end of this step, we obtain the ensemble of proposed models $\mathbf{m}_{prop}^{Q} = [\mathbf{e}_{prop}^{1}, \mathbf{e}_{prop}^{2}, \dots, \mathbf{e}_{prop}^{Q}, \boldsymbol{\pi}_{prop}^{1}, \boldsymbol{\pi}_{prop}^{2}, \dots, \boldsymbol{\pi}_{prop}^{Q}]$.

5) For each chain compute the data generated by the ensemble of perturbed models.

For the *facies move* the prior distribution is used as a proposal distribution, and thus, the acceptance ratio reduces to the ratio of the likelihood probabilities of the configurations $(\boldsymbol{\pi}_{prop}^{q}, \mathbf{m}_{prop}^{q})$ and $(\boldsymbol{\pi}_{curr}^{q}, \mathbf{m}_{curr}^{q})$. Note that the acceptance ratio does not explicitly depend on the facies $\boldsymbol{\pi}$ because the likelihood distribution only depends on $\mathbf{m}$, which is drawn from the prior conditioned by the facies configuration. Therefore, in this case the acceptance probability for each *q*-th chain is equal to:

$$p\left(\mathbf{m}_{prop}^{q} \middle| \mathbf{m}_{curr}^{q}\right) = \min\left[1, \frac{p(\mathbf{d}|\mathbf{m}_{prop}^{q})}{p(\mathbf{d}|\mathbf{m}_{curr}^{q})}\right], \quad (8)$$

Otherwise, the *resistivity move* gives an acceptance probability equal to:

$$p\left(\mathbf{m}_{prop}^{q} \middle| \mathbf{m}_{curr}^{q}\right) = \min\left[1, \frac{p(\mathbf{d}|\mathbf{m}_{prop}^{q})}{p(\mathbf{d}|\mathbf{m}_{curr}^{q})} \frac{p(\mathbf{m}_{prop}^{q})}{p(\mathbf{m}_{curr}^{q})}\right], \quad (9)$$

In this case the proposed configuration $(\boldsymbol{\pi}_{prop}^{q}, \mathbf{m}_{prop}^{q})$ is not directly drawn from the prior and thus, the acceptance probability also depends on the prior ratio term (Gamerman and Lopes, 2006).

6) For each *q*-th chain and only if the perturbed model is accepted set $\mathbf{m}_{curr}^{q} = \mathbf{m}_{prop}^{q}$.

7) Store $\mathbf{m}_{curr}^{q}$ and repeat from 2) to 7) until the maximum number of iterations is reached.

The ensemble of models collected by all the chains after the burn-in is used to compute the statistical properties of the posterior (e.g. mean resistivity model, standard deviation of the resistivity, maximum-a-posteriori facies solution, the posterior probability of occurrence of each facies). All the user-defined parameters (e.g. $\sum_p$, ξ) are properly set to ensure an acceptance ratio during the sampling stage between 0.2-0.4. In particular, ξ should be usually lower than 0.5 because the sampling space of a continuous property is larger than the sampling space of the underlying discrete property (Holmes et al. 2017). In all the following examples we set ξ=0.3. Finally, note that the steps 4) and 5) can be conveniently parallelized. In other terms, the DSS updating step and the forward modeling evaluation for the *Q* chains can be separately run across different CPUs, thus reducing the computing time of the MCMC sampling.

APPLICATION TO SYNTHETIC DATA

In the following synthetic inversion, we assume that the study area is 35 m long and 11 m deep, and this area is discretized with square cells with spatial dimensions of 1 m × 1 m. The schematic synthetic model (Figure 3a) simulates a low resistivity body (50 Ωm) located around 4 m depth and hosted in a high resistivity medium (500 Ωm). Therefore, we employ two different facies: facies A describing the low resistivity medium, and facies B associated with the high resistivity body. Due to the large variability range of the continuous property in the synthetic model, we parametrize the inversion taking the natural logarithm of the resistivity value. For simplicity, we also assume that the prior for the continuous property follows a log-normal Gaussian mixture distribution. In other words, we assume that the natural logarithm of the resistivity values are Gaussian-distributed within each considered facies. However, note that the implemented approach can be also used for non-parametric prior distributions (see the following field data application). Such log-Gaussian mixture prior can be expressed as follows:

$$p(\mathbf{e}) = \sum_{k=1}^{K} \omega_k \mathcal{N}\left(\mathbf{e}; \boldsymbol{\mu}_{\mathbf{e}}^k, \boldsymbol{\Sigma}_{\mathbf{e}}^k\right), \quad (10)$$

where **e** expresses the natural logarithm of the resistivity, *K* is the total number of components of the mixture, ω_k is the prior weight of the *k*-th component, $\mathcal{N}$ indicates the *k*-th Gaussian component with mean vector $\boldsymbol{\mu}_{\mathbf{e}}^{k}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{e}}^{k}$, whereas the superscript *k* indicates that the mean and the covariance are facies dependent. The scalar value ω_k represents the expected frequency of occurrence of the *k*-th facies and for this reason, the *K* values ω_k define the prior model for the discrete property. We assume two different facies hereafter called facies A and facies B: The first is associated with the low resistivity body, while the latter refers to the high resistivity medium. In the following example, we assume that $p(\boldsymbol{\pi} = \text{facies } A) = 0.18$ and $p(\boldsymbol{\pi} = \text{facies } B) = 0.82$. The statistical properties of this prior model can be derived, for example, from borehole data or previous geological information about the study area. We also assume that the spatial variability of the variable **e** follows a stationary (i.e. spatial invariant) 2-D Gaussian variogram model with α_y and α_x equal to 2.5 m and 3.5 m, respectively (see equation 4). The statistical properties of the prior model and the lateral correlation functions associated with the 2-D Gaussian variogram are illustrated in Figure 2.

For simplicity we employ equal lateral and vertical transition matrices that are defined as follows:

$$\mathbf{T}^{l} = \mathbf{T}^{v} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}. \quad (11)$$

For example, if we consider the vertical transition matrix $\mathbf{T}^{v}$ and the *i*-th spatial position, the rows (from top to bottom) represent facies A and facies B at one neighboring position *i*-1, whereas the columns (from left to right) represent facies A and facies B at the *i*-th position. The higher values along the main diagonal preserve the vertical and lateral facies continuity within the sampled models. When dealing with field data inversions these matrices can be inferred from available geological information and expected lateral facies variability (see Vagis et al., 2010 for a detailed discussion on how the transition probabilities can be derived from experimental observations). However, a quality control of the inversion results and slight modifications of the so chosen transition matrices are usually needed to get final predictions with the desired vertical and lateral facies continuity. The previously mentioned FE code has been used to compute the data associated with the schematic

synthetic model. We simulate a Wenner acquisition layout with 36 electrodes evenly spaced of 1 m and an injected current of 1 Ampere. The maximum *a* value we considered is 11. This configuration results in $11 \times 35 = 385$ model parameters to be estimated from 198 data points. Gaussian uncorrelated values with a standard deviation equal to 30 % of the total standard deviation of the noise-free dataset were added to the noise-free pseudosection to simulate noise contamination. This results in a signal-to-noise ratio of 15 dB. The implemented Bayesian approach assumes Gaussian-distributed errors affecting the data. Therefore for simplicity, we are neglecting electrode spacing errors and errors correlated with the magnitude of the observed potential (for more details see Zhou and Dahlin, 2003). However, the next section demonstrates that this Gaussian assumption does not hamper the applicability of the implemented approach to field data.

We run the MCMC inversion using 20 chains evolving for 30000 iterations and the first 5000 models are disregarded from the computation of the PPD to properly burn-in. Figure 3 represents the final results for the continuous property. We observe that the low resistivity anomaly is well recovered, as well as the background resistivity values. As expected, the quality of the predictions decreases at the lateral edges and at the bottom of the model due to the limited illumination. The low spatial resolution of the estimated a-posteriori mean model is related to the employed Gaussian variogram model that smooths the lateral and vertical resistivity contrasts. However, note that the assumed spatial variability patter is crucial to reduce the ill-conditioning of this under-determined inverse problem. The posterior standard deviation map (Figure 3c) shows that the cells located just below the low resistivity body and below 8 m depth are not informed by the data and thus, the associated resistivity predictions are affected by significant uncertainties. Differently, as expected, the lowest posterior uncertainties are located within the rectangular resistivity anomaly and in the shallowest part of the model. These results demonstrate that the implemented approach provides reliable and accurate posterior estimations in agreement with the expected subsurface parameter illumination. Figure 3d shows the evolution of the negative log-likelihood values for the 20 chains

where we observe that the algorithm reaches the stationary regime after 2000-3000 iterations, approximately.

Figure 4 shows the results for the discrete property. The spatial location of the low resistivity body is well recovered by the maximum-a-posterior (MAP) facies solution (Figure 4b). As expected the shallowest edge of the low resistivity body is perfectly recovered, while the accuracy of the MAP solution decreases as the depth increases due to the decreased parameter illumination. The posterior probability for the discrete property (Figures 4c and 4d) shows that the occurrence of the two facies is well recovered in the shallowest part of the model and within the low resistivity body. Conversely, the uncertainty increases as the depth increases due to the reduced illumination. Again, the low spatial resolution of the probability values shown in Figures 4c and 4d are related to the Gaussian variogram model that acts as a regularization operator in the MCMC inversion. Table 1 shows the percentage of data and model errors computed at the first and last MCMC iterations. The data error was calculated as the difference between the observed data and the data computed on the mean resistivity models at the first and last iterations. Similarly, the model error was calculated as the difference between the true resistivity model and the mean models predicted at the first and last iterations.

In Figure 5 we represent some examples of marginal posterior distributions for the discrete and continuous property estimated at three different spatial locations. As expected, the posterior model accurately predicts the actual facies configuration and resistivity values in the shallowest part of the model and within the low resistivity anomaly (Figures 5a and 5b). In particular, note that in these cases the MAP solution for the continuous property is very close to the true resistivity value. Conversely, Figure 5c shows that the posterior model is very similar to the prior for the deepest cells, thereby demonstrating that this part of the subsurface is not informed by the data.

In Figure 6a we represent a direct comparison between the observed data (i.e. the apparent resistivity pseudosection flatted to a 1-D vector) and the data associated with the current models at different MCMC iterations. We observe that the observed data are well predicted after 5000 iterations, thus demonstrating that the stationary regime is attained. Figures 6b-c compares the observed and the

predicted pseudosections. Again we observe a good match between the observed and predicted datasets.

Figure 7 shows 12 out of 20 facies models and associated resistivity values extracted from the prior distribution $p(\mathbf{m})$ that constitute some examples of starting points for the MCMC inversion. Figure 8 shows 12 out of 20 models sampled at the final MCMC iteration that can be considered as an ensemble of facies and resistivity configurations drawn from the equivalence region of solutions in the parameter space (Fernandez Martinez et al., 2011). In other terms, all these models generate predicted data that equally reproduce the observed data. Note that a low resistivity anomaly is clearly visible at the center of these models, while their differences increase as the depth increases, that is as the parameter illumination decreases. Finally, Figure 9 shows that for all the 385 continuous model parameters the MCMC inversion has reached a stable posterior model as demonstrated by PSRF values lower than the threshold of 1.2. In particular, note that the convergence to a stable PPD for the vast majority of the parameters is reached within 20000 iterations, then before the maximum allowable number of iterations.

APPLICATION TO FIELD DATA

We applied the implemented MCMC inversion to 2D field data acquired by a permanent monitoring system installed along a river embankment for monitoring the inner conditions of the levee. Besides reducing the hydrogeological risks, long-term ERT monitoring data can be used to understand saturation and dewatering of levee structures, and to analyze how resistivity values are related to annual variations of the water content (Tresoldi et al., 2019). In this work, we limit to invert a single dataset and we refer the reader to Hojat et al. (2019b) for more information about the study site. The electrode layout is buried in a 0.5 m-deep trench. We used the data corrected for the effect of the soil overlaying electrodes (Hojat et al. 2019c; Hojat et al. 2020). The inversion covers an area that is 94 m wide and 14 m deep and has been discretized with rectangular cells with dimensions of 1 m and 2 m along the vertical and horizontal directions, respectively. The acquisition layout

comprises a Wenner acquisition geometry with 48 electrodes and the unit electrode spacing a = 2 m. This configuration results in 705 resistivity values to be estimated from 360 data points. Note that the forward modeling code embeds the inverted domain into a larger domain to avoid boundary effects and numerical artifacts (Maurer and Friedel, 2006).

To define the a-priori facies and resistivity distribution we exploited both the available geological information about the investigated area and the multiple data and associated predicted resistivity sections obtained during the permanent monitoring. In particular, we expect a mainly clay body hosting a more permeable layer constituted by sand and gravel located at around 2-3 m depth. Therefore, we assume two different facies (i.e. facies A and B) associated with the clay and sand formations and characterized by low and high resistivity values, respectively. The prior facies probabilities are equal to $p(\boldsymbol{\pi} = \text{facies } A) = 0.85$ and $p(\boldsymbol{\pi} = \text{facies } B) = 0.15$, while the kernel density estimation algorithm has been used to numerically compute the non-parametric prior distribution of the resistivity values within each facies. The final non-parametric marginal prior model for the continuous property can be written as follows:

$$p(\mathbf{e}) = \sum_{k=1}^{K} \omega_k \, p(\mathbf{e}|\boldsymbol{\pi}^k), \quad (12)$$

where K=2, whereas $\omega_1 = 0.85$, and $\omega_2 = 0.15$. Figures 10a and 10b represent the prior model for both the discrete and the continuous property. Repeated measures of the apparent resistivity values taken within a short time frame give us an estimate of the uncertainty affecting the observed data that is used to build the data covariance matrix. We assume a 2-D stationary Gaussian variogram model with α_y and α_x values equal to 2 m and 6 m, respectively. The lateral and vertical correlation functions associated with this variogram model are represented in Figures 10c and 10d. Similar to the synthetic inversion we employ equal lateral and vertical transition matrices that are defined as follows:

$$\mathbf{T}^l = \mathbf{T}^v = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}. \quad (13)$$

In this application, the extensive geological information available helped us in the definition of an appropriate prior model for the study area. However, the additional tests we carried out demonstrated that the final results do not significantly change if we use, for example, the same transition matrix employed in the synthetic example. This is expected because in the case of high-quality data and for the most illuminated parameters, the final results are mainly driven by the data likelihood function instead of the prior assumptions. The FE code previously used in the synthetic inversion again constitutes the forward modeling engine for the inversion process. Similarly to the synthetic example, we employ 20 chains running for 30000 iterations.

Figure 11a shows the result provided by a local least-squares inversion approach (Loke, 2018), whereas Figure 11b illustrates the posterior mean model estimated by the implemented MCMC algorithm. The similarity of the two inversion outcomes ensures us about the reliability of the final solution and also confirms the suitability of the proposed MCMC algorithm for real data applications. As expected from previous inversion results obtained in the same area and from the available geological information, both algorithms predict a high resistivity body around 2 m depth (associated with sand/gravel) hosted in a low resistivity medium (clay). Differently from the local inversion, the MCMC also provides an accurate estimate of the uncertainty affecting the retrieved solution that is represented in Figure 11c in the form of a standard deviation section. As expected the shallowest part of the subsurface is characterized by the lowest uncertainty while the ambiguity increases within the high resistivity formation, at the lateral edges and at the deepest part of the model due to the lower illumination.

Figure 12 shows the MAP facies solution and the marginal posterior probability for the discrete property estimated by the MCMC inversion overt the study area. We observe that the high resistivity formation has the highest probability of occurrence around 2 m depth. Figure 13 compares the observed data and the two pseudosections computed from the model estimated by the local inversion and from the mean posterior solution shown in Figure 11b. We observe that both inversion approaches provide similar predicted data characterized by a good match with the measured pseudosection. More

quantitatively, the final percentage data errors for the gradient-based and MCMC approachs are 1.35 % and 1.18 %, respectively. The slightly higher misfit provided by the MCMC inversion can be attributed to the many model constraints (i.e. for both the discrete and the contoìinuos property) infused into the inversion framework.

The evolution of the negative log-likelihood for the 20 chains highlights the fast convergence of the MCMC sampling toward the stationary regime that is attained after 2000-3000 iterations (Figure 14a). The evolution of the PSRF confirms (Figure 14b) that the convergence to a stable posterior model for all the continuous model parameters has been achieved within the maximum number of iterations allowed. In particular, note that only 20000 iterations are needed to reach a stable PPD for the vast majority of the unknowns.

Figure 15 shows examples of marginal prior and posterior probabilities computed for some cells. In Figures 15a and 15b we observe that in the shallow part of the model the PPDs are very different from the prior thereby confirming that these cells are informed by the data. We can also appreciate how the posterior uncertainty substantially increases moving from the low resistivity body located in the shallowest part of the model to the high resistivity formation located around 2 m depth. As expected the prior and the posterior are very similar for the cell located at the lateral edge of the model where the data illumination is poor (Figure 15c).

Finally, Figure 16 illustrates some examples of facies and corresponding resistivity models extracted from the a-priori assumptions and forming the starting points of the MCMC inversion, wheres Figure 17 shows some facies configurations and associated resistivity values sampled at the end of the MCMC inversion. Note in Figure 17 that all the considered chains univocally predict a high resistivity body located around 2 m depth.

DISCUSSION

The implemented Markov Chain Monte Carlo (MCMC) inversion is aimed at accurately assessing the multi-modal posterior distribution of facies and resistivity values in a high-dimensional parameter

space. In our application, the DEMC approach has been used to sample the posterior model. Before selecting this MCMC method we carried out many inversion tests in which we compared the sampling abilities of different MCMC algorithms (i.e. Random Walk Metropolis, Adaptive Metropolis, Metropolis algorithm with adaptation of the scaling factor, Differential Evolution Markov Chain, and the Differential Evolution Adaptive Metropolis) on analytical probability density functions and on synthetic geophysical inversion problems (e.g., cross-hole traveltime tomography, ERT inversion). These tests proved thatthe DEMC constitutes the optimal compromise between the converge rate toward a stable posterior and the computational effort in all examples.

In our inversion examples, we assume the noise to be uncorrelated and Gaussian-distributed. Possible improvements could be including the noise spatial correlation and/or the noise standard deviation as additional unknowns to be sampled during the inversion (Bodin et al. 2012). From a mathematical point of view, the 2-D implemented approach can be easily extended to 3-D models by adding additional lateral constrains for the continuous and discrete model parameters. One of the challenges of 3-D applications is related to the curse of dimensionality problem (Curtis and Lomax 2001) that can drastically decrease the convergence of the algorithm to a stationary posterior model because in high-dimensional problems the PPD tends to be highly localized within each model dimension. This results in a low acceptance ratio and a very slow convergence of the probabilistic sampling. In other words, the chain tends to converge to local optima of the PPD associated with a given facies configuration. To partially overcome this situation, we suggest using a high ξ value (around 0.8-0.9) at the beginning of the inversion thus promoting jumps between different facies configurations. Then, the ξ value should be decreased (i.e. to 0.2-0.3) during the sampling to get a reliable estimate of the posterior probability of the continuous property.

There are usually no rigid rules to set the MCMC hyperparameters. In our application, we set the user-defined parameters $(\xi, \sum_p, \gamma)$ through a trial and error procedure in which we monitored the acceptance ratio during the sampling stage. Indeed, it is known that an acceptance ratio lying in the range [0.2, 0.4] usually offers the best compromise between the exploitation and the exploration of

the algorithm. In our experiments, we also found that ξ and γ are the parameters that mostly influence the acceptance ratio value. It is also known that the use of the a-priori as the proposal usually reduces the acceptance ratio of the chain and slows down the convergence. For this reason, the algorithm can be improved by sampling from an approximate posterior distribution rather than from the prior assumption. To this end many strategies could be adopted, for example, sampling from an approximated estimate of the posterior covariance matrix computed from a local approximation of the Jacobian matrix (see Dosso et al. 2014). Another viable strategy to reduce the burn-in period could be starting the MCMC sampling from the model predicted by a local inversion. More advanced MCMC algorithms that incorporate the principles of Hamiltonian dynamics into the standard Metropolis-Hasting method (Betancourt, 2017; Fichtner et al. 2019; Gebraad et al. 2020; Aleardi and Salusti 2020b; Aleardi et al. 2020) could be useful to speed up the probabilistic ERT inversion. The major computational requirement of the Hamiltonian Monte Carlo algorithm is the need for computing the derivative (i.e. the Jacobian matrix) for each sampled model. To this end, the adjoint-state method can be employed. Alternatively, one can also exploit Machine learning algorithms to approximate the Jacobian around each considered model (Aleardi, 2020).

The MCMC inversion and the employed FE forward modeling code have been implemented in Matlab. In particular, we used a parallel MCMC code in which the model update and the forward evaluations are distributed across different cores. The synthetic inversion and real data examples run in approximately 1.5 and 2 days, approximately, on a two deca-core intel E5-2630 at 2.2 GHz (128 Gb RAM). The major computational requirement, as expected, is related to the forward modeling evaluation that covers more than the 85% of the entire computational cost of the implemented inversion approach. Therefore, the MCMC inversion is much more computationally demanding than a gradient-based inversion that runs in a few minutes on the same hardware. Although this is an important disadvantage to invert monitoring data that require real-time results (Hojat et al., 2020), the method can be efficiently used to get reasonable results for a variety of studies (e.g. archaeological studies, mineral exploration, groundwater studies). The main advantage of the MCMC inversion over

the local approach is the possibility to reliably assess the uncertainty affecting the recovered solution that can be used to build multiple realizations of both the continuous and discrete properties. Such realizations can be extremely useful during the geological interpretation phase because represent possible subsurface scenarios all in accordance with the recorded data.

If needed, the use of geostatistical simulation approaches allows for a direct inclusion of hard model constraints (e.g. facies or resistivity constraints provided by borehole data) into the implemented inversion framework. Finally, we point out that there is still room for a substantial reduction of the computational cost for example by employing a more efficient and scalable forward modeling code. For this reason, we are exploring the possibility to translate the Matlab code into a low-level programming language.

CONCLUSIONS

We implemented an MCMC inversion algorithm for Bayesian ERT inversion in which the subsurface resistivity model, the facies configuration, and the associated uncertainties are inferred from the measured apparent resistivity values. The proposed approach provides accurate uncertainty estimations in case of mixed discrete-continuous inverse problems, non-linear forward models, and multimodal, non-parametric prior distributions. In other terms, our approach does not require any assumptions (i.e. Gaussianity) about the distribution of the continuous model property in a given facies. To reduce the ill-conditioning of the inversion process we included spatial constraints for both the continuous and discrete model parameters. In particular, we used a Gaussian variogram model for the resistivity value, and a 1D Markov prior models for the facies, whereas a FE code was used for forward modeling. The synthetic and field inversions showed that the implemented MCMC approach yielded reliable final predictions in accordance with the outcomes of a gradient-based inversion. However, the outstanding benefit of the MCMC algorithm is the possibility to accurately assess the uncertainties affecting the recovered solution.

ACKNOWLEDGMENTS

This research was partially funded by Ministero dell'Ambiente e della Tutela del Territorio e del Mare, Italy, project DILEMMA – Imaging, Modeling, Monitoring and Design of Earthen Levees. We are grateful to LSI-Lastem S.r.l. that designed and installed the G.RE.T.A. (Geo REsistivimeter for Time-lapse Analysis) prototype to monitor the Colorno pilot site.

Conflict of interest

The authors declare no conflict of interest

Data Availability Statement

Data available on request from the authors.

REFERENCES

Aleardi, M., Ciabarri, F., and Gukov, T. (2018). A two-step inversion approach for seismic-reservoir characterization and a comparison with a single-loop Markov-chain Monte Carlo algorithm. Geophysics, 83(3), R227-R244.

Aleardi, M., and Salusti, A. (2020a). Markov chain Monte Carlo algorithms for target-oriented and interval-oriented amplitude versus angle inversions with non-parametric priors and non-linear forward modellings. Geophysical Prospecting, 68(3), 735-760.

Aleardi, M., and Salusti, A. (2020b). Hamiltonian Monte Carlo algorithms for target-and interval-oriented amplitude versus angle inversions. Geophysics, 85(3), R177-R194.

Aleardi, M., Salusti, A., and Pierini, S. (2020). Transdimensional and Hamiltonian Monte Carlo inversions of Rayleigh-wave dispersion curves: A comparison on synthetic datasets. Near Surface Geophysics. In print. doi: 10.1002/nsg.12100.

Aleardi, M. (2020). Combining discrete cosine transform and convolutional neural networks to speed up the Hamiltonian Monte Carlo inversion of pre-stack seismic data. Geophysical Prospecting. In print. doi: 10.1111/1365-2478.13025.

Aleardi, M., and Mazzotti, A. (2017). 1D elastic full-waveform inversion and uncertainty estimation by means of a hybrid genetic algorithm–Gibbs sampler approach. Geophysical Prospecting, 65(1), 64-85.

Aleardi, M., Ciabarri, F., and Mazzotti, A. (2017). Probabilistic estimation of reservoir properties by means of wide-angle AVA inversion and a petrophysical reformulation of the Zoeppritz equations. Journal of Applied Geophysics, 147, 28-41.

Andersen, K. R., Wan, L., Grombacher, D., Lin, T., and Auken, E. (2018). Studies of parameter correlations in surface NMR using the Markov chain Monte Carlo method. Near Surface Geophysics, 16(2), 206-217.

Arosio, D., Munda, S., Tresoldi, G., Papini, M., Longoni, L., and Zanzi, L. (2017). A customized resistivity system for monitoring saturation and seepage in earthen levees: installation and validation. Open Geosciences, 9(1), 457-467. DOI: 10.1515/geo-2017-0035.

Azevedo, L., and Soares, A. (2017). Geostatistical methods for reservoir geophysics. Springer, Switzerland.

Betancourt, M. (2017). A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint arXiv:1701.02434.

Bodin, T., Sambridge, M., Tkalčić, H., Arroucau, P., Gallagher, K., and Rawlinson, N. (2012). Transdimensional inversion of receiver functions and surface wave dispersion. Journal of Geophysical Research: Solid Earth, 117, B02301.

Crawford, M. M., Bryson, L. S., Woolery, E. W., and Wang, Z. (2018). Using 2-D Electrical resistivity imaging for joint geophysical and geotechnical characterization of shallow landslides. Journal of Applied Geophysics, 157, 37-46.

Curtis, A., and Lomax, A. (2001). Prior information, sampling distributions, and the curse of dimensionality. Geophysics, 66(2), 372-378.

Dosso, S. E., Dettmer, J., Steininger, G., and Holland, C. W. (2014). Efficient trans-dimensional Bayesian inversion for geoacoustic profile estimation. Inverse Problems, 30(11), 114018.

Doyen, P. (2007). Seismic reservoir characterization: An earth modelling perspective. EAGE publications, Houten.

Fernández-Martínez, J. L., Mukerji, T., García-Gonzalo, E., and Fernández-Muñiz, Z. (2011). Uncertainty assessment for inverse problems in high dimensional spaces using particle swarm optimization and model reduction techniques. Mathematical and Computer Modelling, 54(11-12), 2889-2899.

Fiandaca, G., Madsen, L. M., and Maurya, P. K. (2018). Re-parameterisations of the Cole–Cole model for improved spectral inversion of induced polarization data. *Near Surface Geophysics*, *16*(4), 385-399.

Fichtner, A., Zunino, A., and Gebraad, L. (2019). Hamiltonian Monte Carlo solution of tomographic inverse problems. Geophysical Journal International, 216(2), 1344-1363.

Galetti, E., and Curtis, A. (2018). Transdimensional Electrical resistivity tomography. Journal of Geophysical Research: Solid Earth, 123(8), 6347-6377.

Gamerman, D., and Lopes, H. F. (2006). Markov chain Monte Carlo: stochastic simulation for Bayesian inference. CRC Press.

Gebraad, L., Boehm, C., and Fichtner, A. (2020). Bayesian elastic Full-Waveform Inversion using Hamiltonian Monte Carlo. Journal of Geophysical Research: Solid Earth, 125(3), e2019JB018428

Gelman, A., Carlin, J.B., Stern, H.S. and Rubin, D.B. (1995). Bayesian Data Analysis, Chapman & Hall, London.

Grana, D., and Della Rossa, E. (2010). Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion. Geophysics, 75(3), O21-O37.

Gunning, J., and Sams, M. (2018). Joint facies and rock properties Bayesian amplitude-versus-offset inversion using Markov random fields. Geophysical Prospecting, 66(5), 904-919.

Hermans, T., and Paepen, M. (2020). Combined inversion of land and marine Electrical resistivity tomography for submarine groundwater discharge and saltwater intrusion characterization. Geophysical Research Letters, 47(3), e2019GL085877.

Hojat, A., Arosio, D., Ivanov, V. I., Longoni, L., Papini, M., Scaioni, M., Tresoldi, G., and Zanzi, L. (2019a). GeoElectrical characterization and monitoring of slopes on a rainfall-triggered landslide simulator. Journal of Applied Geophysics, 170, 103844. DOI: 10.1016/j.jappgeo.2019.103844.

Hojat, A., Arosio, D., Longoni, L., Papini, M., Tresoldi, G., and Zanzi, L. (2019b). Installation and validation of a customized resistivity system for permanent monitoring of a river embankment. EAGE-GSM 2nd Asia Pacific Meeting on Near Surface Geoscience and Engineering, Kuala Lumpur, Malaysia, DOI: 10.3997/2214-4609.201900421.

Hojat A., Arosio D., Loke M.H., Longoni L., Papini M., Tresoldi G. and Zanzi L. (2019c). Assessment of 3D geometry effects on 2D ERT data of a permanent monitoring system along a river embankment. EAGE-GSM 2nd Asia Pacific Meeting on Near Surface Geoscience and Engineering, Kuala Lumpur, Malaysia, DOI: 10.3997/2214-4609.201900427.

Hojat, A., Arosio, D., Ivanov, V.I., Loke, M.H., Longoni, L., Papini, M., Tresoldi, G. and Zanzi, L. (2020) Quantifying seasonal 3D effects for a permanent electrical resistivity tomography (ERT) monitoring system along the embankment of an irrigation canal. Near Surface Geophysic 18, 427-443, DOI: 10.1002/nsg.12110.

Holmes, C., Krzysztof, L. and Pompe, E. (2017). Adaptive MCMC for multimodal distributions. Technical report. https://pdfs.semanticscholar.org/c75d/f035c23e3c0425409e70d457cd43b174076f.pdf.

Horta, A., and Soares, A. (2010). Direct sequential co-simulation with joint probability distributions. Mathematical Geosciences, 42(3), 269-292.

Irving, J., and Singha, K. (2010). Stochastic inversion of tracer test and electrical geophysical data to estimate hydraulic conductivities. Water Resources Research, 46, W11514.

Jardani, A., Revil, A., and Dupont, J. P. (2013). Stochastic joint inversion of hydrogeophysical data for salt tracer test monitoring and hydraulic conductivity imaging. Advances in Water Resources, 52, 62–77.

Johnson, T. C., Versteeg, R. J., Ward, A., Day-Lewis, F. D., and Revil, A. (2010). Improved hydrogeophysical characterization and monitoring through parallel modeling and inversion of time-domain resistivity and induced-polarization data. Geophysics, 75(4), WA27-WA41.

Kaipio, J. P., Kolehmainen, V., Somersalo, E., and Vauhkonen, M. (2000). Statistical inversion and Monte Carlo sampling methods in electrical impedance tomography. Inverse Problems, 16(5), 1487.

Karaoulis, M., Revil, A., Tsourlos, P., Werkema, D. D., and Minsley, B. J. (2013). IP4DI: A software for time-lapse 2D/3D DC-resistivity and induced polarization tomography. Computers & Geosciences, 54, 164-170.

Karaoulis, M., Tsourlos, P., Kim, J. H., and Revil, A. (2014). 4D time-lapse ERT inversion: introducing combined time and space constraints. Near Surface Geophysics, 12(1), 25-34.

Legaz, A., Vandemeulebrouck, J., Revil, A., Kemna, A., Hurst, A. W., Reeves, R., and Papasin, R. (2009). A case study of resistivity and self-potential signatures of hydrothermal instabilities, Inferno Crater Lake, Waimangu, New Zealand. Geophysical Research Letters, 36(12).

Linde, N., Binley, A., Tryggvason, A., Pedersen, L. B., and 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.

Loke, M. H. (2018). Rapid 2-D Resistivity and IP inversion using the least-squares method. www.geotomosoft.com, last accessed on 3 July 2020.

Maurer, H., and Friedel, S. (2006). Outer-space sensitivities in geoelectrical tomography. Geophysics, 71(3), G93-G96.

Moradipour, M., Ranjbar, H., Hojat, A., Karimi-Nasab, S. and Daneshpajouh, S. (2016) Laboratory and field measurements of electrical resistivity to study heap leaching pad no. 3 at

Sarcheshmeh copper mine. 22nd European Meeting of Environmental and Engineering Geophysics, DOI: 10.3997/2214-4609.201602140.

Müller, K., Vanderborght, J., Englert, A., Kemna, A., Huisman, J. A., Rings, J., and Vereecken, H. (2010). Imaging and characterization of solute transport during two tracer tests in a shallow aquifer using Electrical resistivity tomography and multilevel groundwater samplers. Water Resources Research, 46(3).

Pidlisecky, A., and Knight, R. (2008). FW2_5D: A MATLAB 2.5-D Electrical resistivity modeling code. Computers & Geosciences, 34(12), 1645-1654.

Pollock, D., and Cirpka, O. A. (2012). Fully coupled hydrogeophysical inversion of a laboratory salt tracer experiment monitored by Electrical resistivity tomography. Water Resources Research, 48(1).

Ramirez, A. L., Nitao, J. J., Hanley, W. G., Aines, R., Glaser, R. E., Sengupta, S. K., Dyer, K.M, Hickling, T.L., and Daily, W. D. (2005). Stochastic inversion of Electrical resistivity changes using a Markov Chain Monte Carlo approach. Journal of Geophysical Research: Solid Earth, 110(B2).

Sambridge, M., and Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. Reviews of Geophysics, 40(3), 3-1.

Scalzo, R., Kohn, D., Olierook, H., Houseman, G., Chandra, R., Girolami, M., and Cripps, S. (2019). Efficiency and robustness in Monte Carlo sampling for 3-D geophysical inversions with Obsidian v0. 1.2: setting up for success. Geoscientific Model Development, 12(7), 2941-2960.

Sen, M. K., and Stoffa, P. L. (2013). Global optimization methods in geophysical inversion. Cambridge University Press.

Siemon, B., van Baaren, E., Dabekaussen, W., Delsman, J., Dubelaar, W., Karaoulis, M., and Steuer, A. (2019). Automatic identification of fresh–saline groundwater interfaces from airborne electromagnetic data in Zeeland, the Netherlands. Near Surface Geophysics, 17(1), 3-25.

Soares, A. (2001). Direct sequential simulation and cosimulation. Mathematical Geology, 33(8), 911-926.

Tarantola, A. (2005). Inverse problem theory, siam.

Ter Braak, C. J. (2006). A Markov Chain Monte Carlo version of the genetic algorithm Differential Evolution: easy Bayesian computing for real parameter spaces. Statistics and Computing, 16(3), 239-249.

Tresoldi, G., Arosio, D., Hojat, A., Longoni, L., Papini, M., and Zanzi, L. (2019). Long-term hydrogeophysical monitoring of the internal conditions of river levees. Engineering Geology, 259, 105139. DOI: 10.1016/j.enggeo.2019.05.016.

Turner, B. M., Sederberg, P. B., Brown, S. D., and Steyvers, M. (2013). A method for efficiently sampling from distributions with correlated dimensions. Psychological methods, 18(3), 368.

Ulvmoen, M., and Omre, H. (2010). Improved resolution in Bayesian lithology/fluid inversion from prestack seismic data and well observations: Part 1—Methodology. Geophysics, 75(2), R21-R35.

Vagis, A. A. (2010). Bayesian Recognition Procedures on Networks. Journal of Automation and Information Sciences, 42(11), 58-63.

Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. Environmental Modelling & Software, 75, 273-316.

Whiteley, J., Chambers, J.E., and Uhlemann, S., 2017. Integrated monitoring of an active landslide in lias group mudrocks, north yorkshire, UK. In: Hoyer, S. (Ed.), GELMON 2017: 4th International Workshop on GeoElectrical Monitoring, 27.

Zhang, Y., Ghodrati, A., and Brooks, D. H. (2005). An analytical comparison of three spatio-temporal regularization methods for dynamic linear inverse problems in a common statistical framework. Inverse Problems, 21(1), 357.

Zhou, B., and Dahlin, T. (2003). Properties and effects of measurement errors on 2D resistivity imaging surveying. Near surface geophysics, 1(3), 105-117.

FIGURES

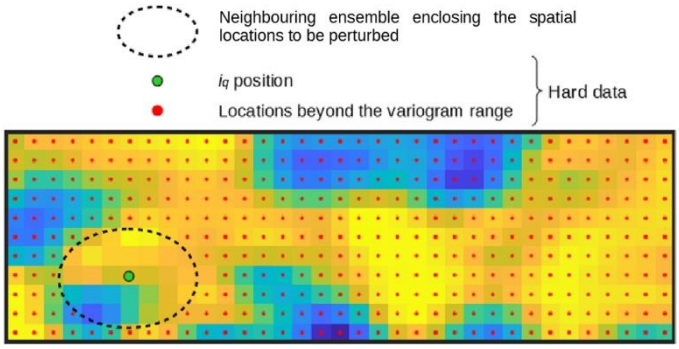

Figure 1: Example of a resistivity model in which we highlight the hard data values and the neighboring locations where the current resistivity values will be modified by the DSS updating step.

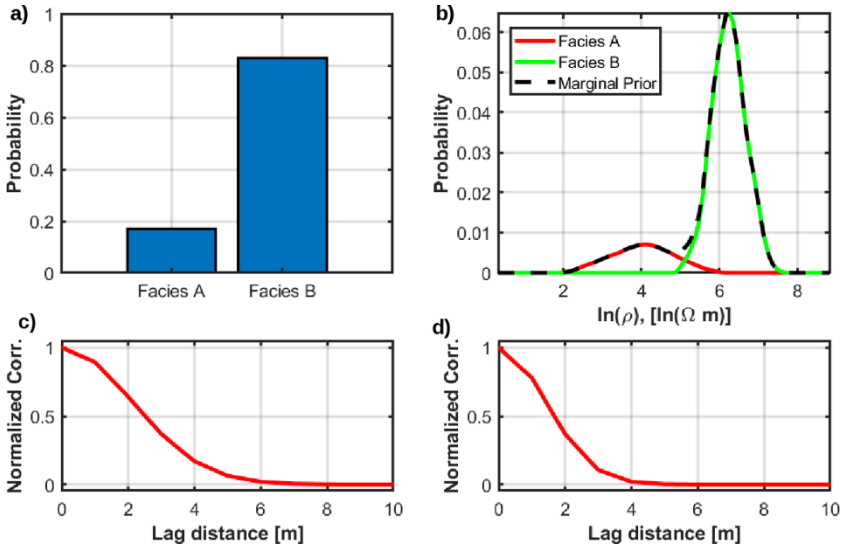

Figure 2: a) Prior distribution for the discrete property showing the expected frequency of occurrence of the two facies in the synthetic example. b) Marginal prior distribution for the natural logarithm of the resistivity values. The red and green curves show the Gaussian components forming the Gaussian mixture model (black curve). c), and d) spatial correlation functions associated with the assumed 2-D variogram model along the lateral and vertical directions, respectively.

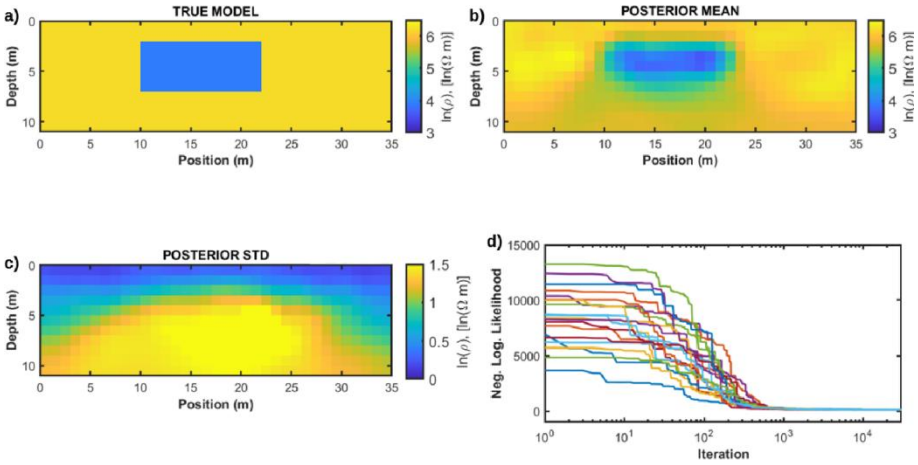

Figure 3: a) The true model for the synthetic inversion. b) The a-posteriori mean solution estimated by the MCMC inversion. c) The standard deviation estimated by the MCMC inversion. d) Evolution of the negative log-likelihood value for the 20 chains.

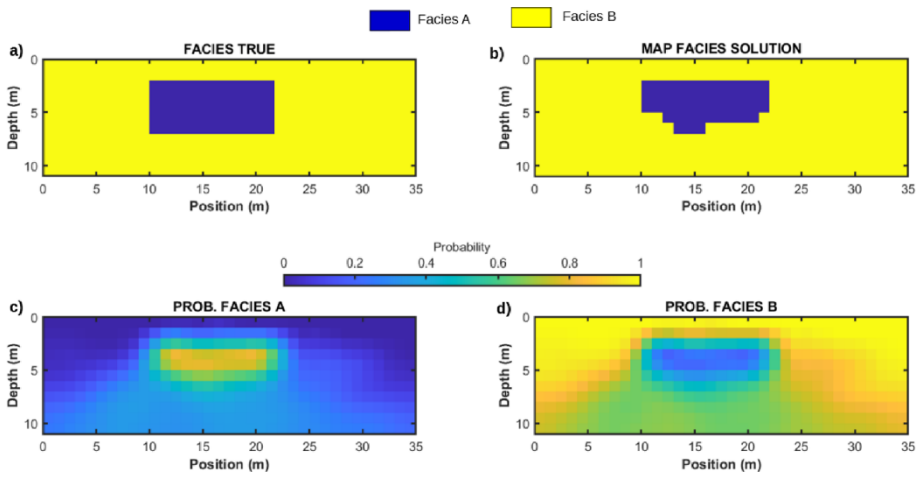

Figure 4: a) True facies model. b) Maximum-a-posteriori (MAP) facies solution. c), and d) posterior probabilities for the discrete property showing the probability of observing the two facies at each subsurface spatial location.

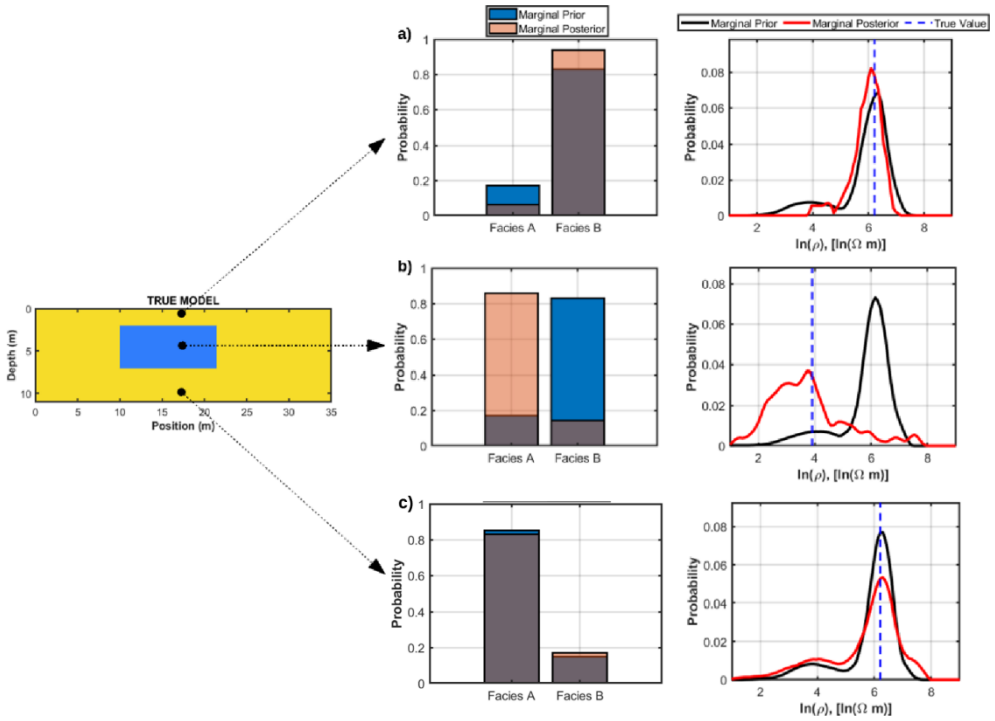

Figure 5: a), b), and c) examples of marginal posterior distributions for the discrete and continuous property at three different spatial locations.

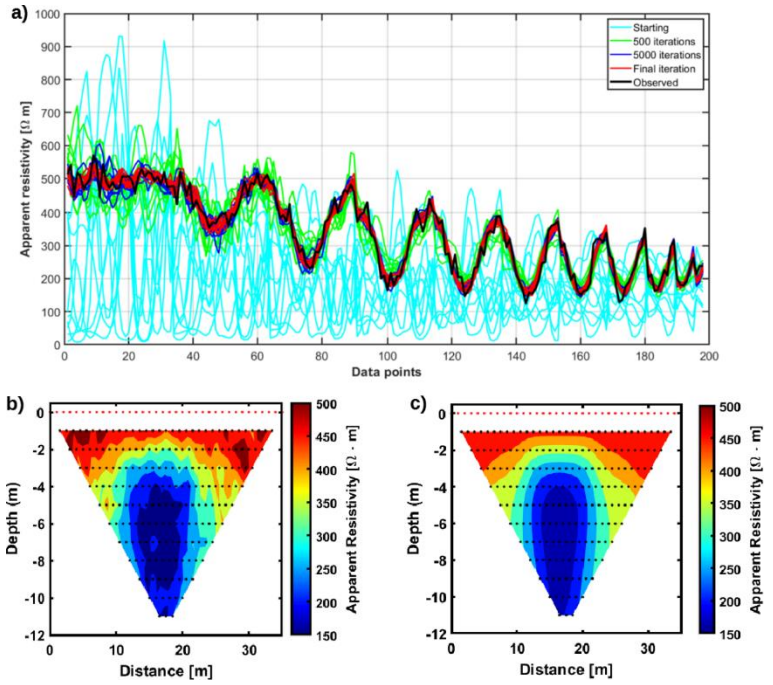

Figure 6: a) Comparison between the observed data and the data predicted at different MCMC iterations. For graphical convenience, the apparent resistivity pseudosections have been flattened to 1-D vectors. Note that 5000 iterations are enough to accurately predict the observed data. b) The observed pseudosection. c) The pseudosection computed on the a-posterior mean model (Figure 3b).

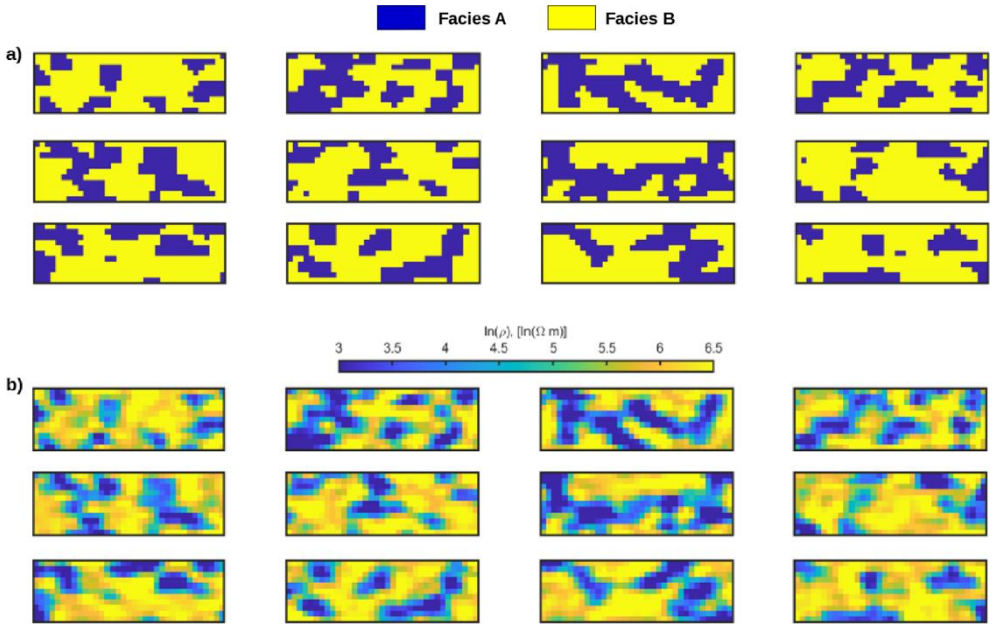

Figure 7: Examples of facies configurations (a) and associated resistivity values (b) extracted from the a-priori distribution and forming 12 out of the 20 current models of the first MCMC iteration.

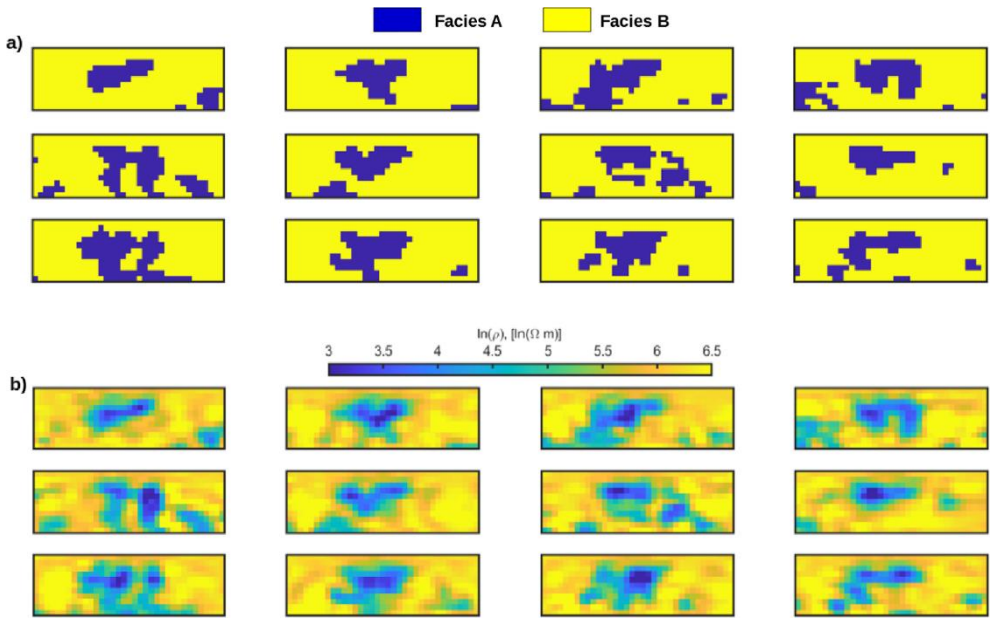

Figure 8: Examples of facies configurations (a) and associated resistivity values (b) forming 12 out of the 20 current models of the last MCMC iteration.

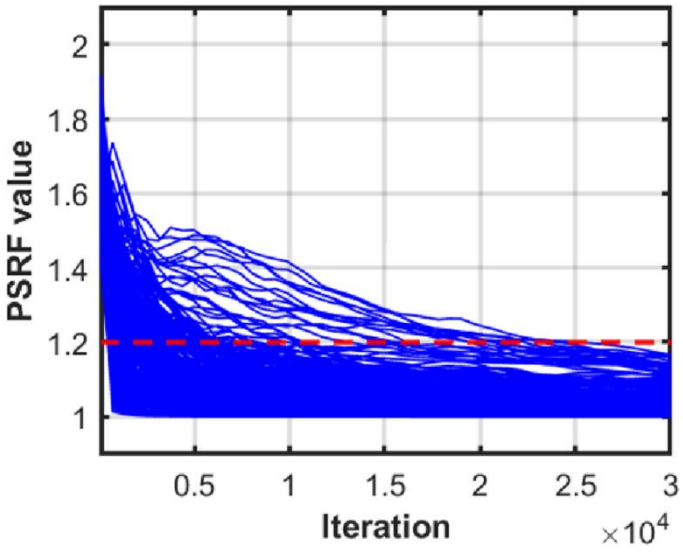

Figure 9: Evolution of the PSRF values for all the 385 model parameters. The dotted red line at 1.2 indicates the threshold of convergence.

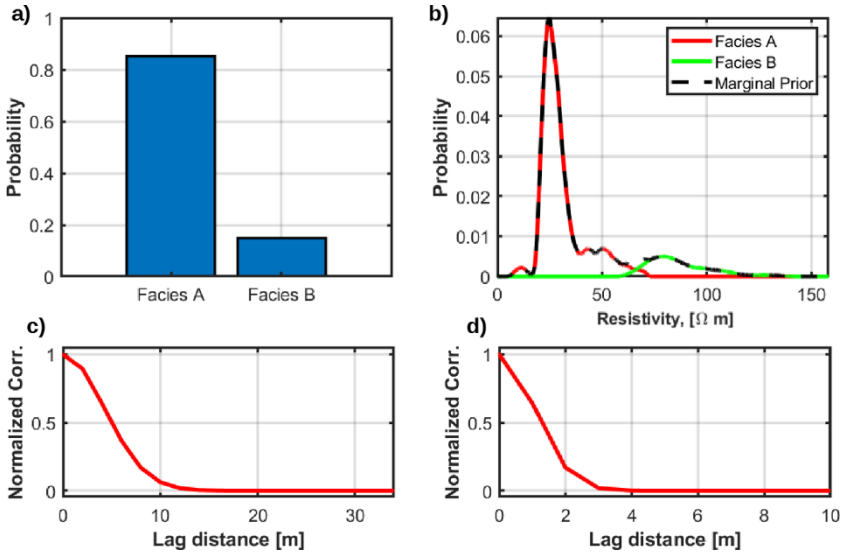

Figure 10: a) Prior distribution for the discrete property showing the expected frequency of occurrence of the two facies in the investigated area. Facies A and B refer to the low and high resistivity formations. b) Non-parametric marginal prior distribution for the resistivity values (dotted black curve). The red and green curves show the non-parametric components associated with each facies. c), and d) normalized spatial correlation functions associated with the assumed 2D Gaussian variogram model along the lateral and vertical directions, respectively.

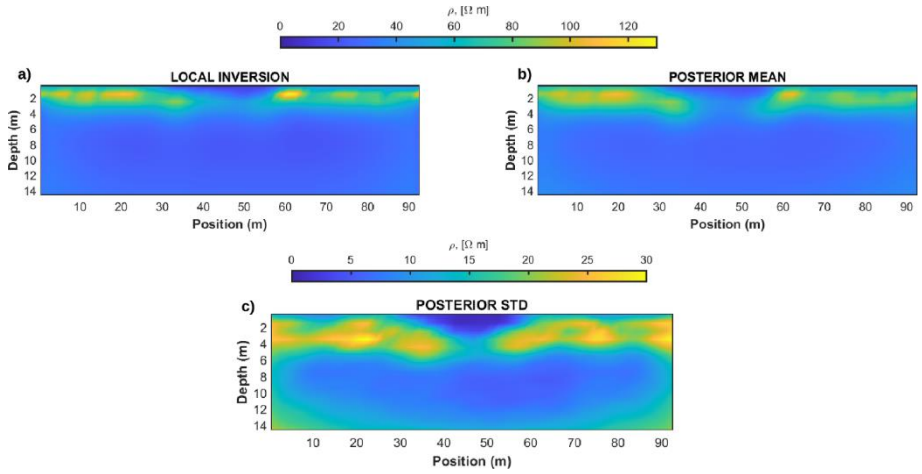

Figure 11: a) The model estimated by a gradient-based least-squares inversion. b) The a-posteriori mean solution estimated by the MCMC inversion. c) The standard deviation map estimated by the MCMC inversion.

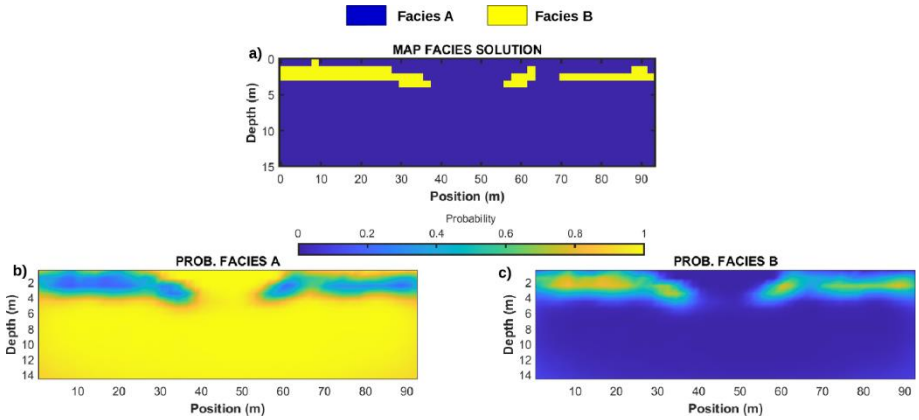

Figure 12: a) Maximum-a-posteriori (MAP) facies solution. b), and c) posterior probabilities for the discrete property showing the probability of observing the two facies at each subsurface spatial location.

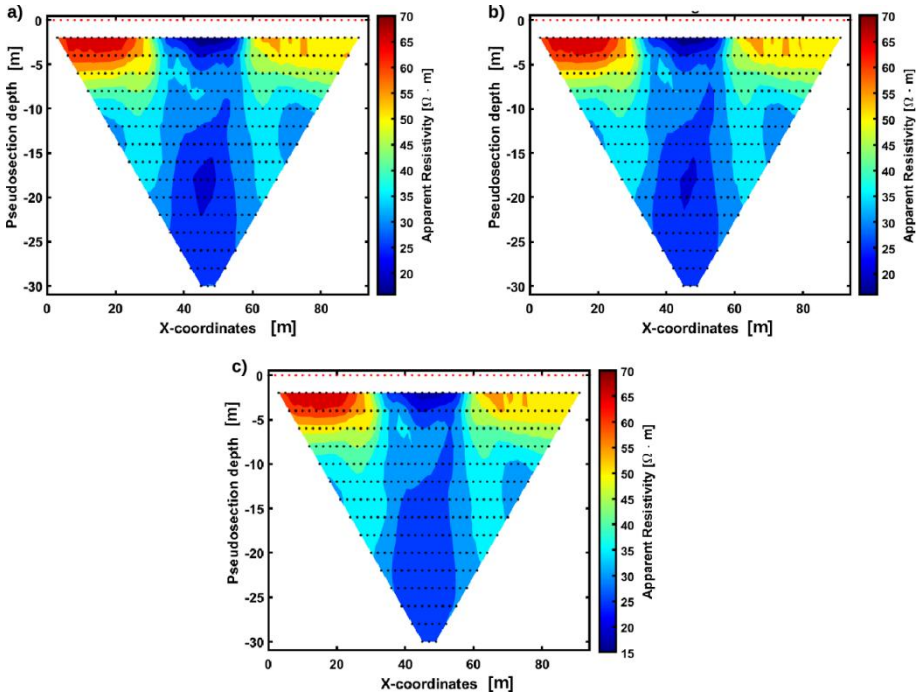

Figure 13: a) Observed pseudo-section. b) Pseudo-section predicted by a local least-squares inversion. c) Pseudo-section computed on the mean posterior solution provided by the MCMC inversion.

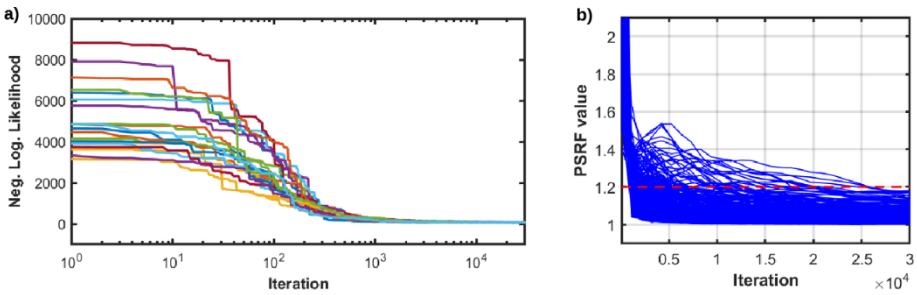

Figure 14: a) Evolution of the negative log-likelihood value for the 20 chains. b) Evolution of the PSRF values for all the 705 model parameters. The dotted red line at 1.2 indicates the threshold of convergence.

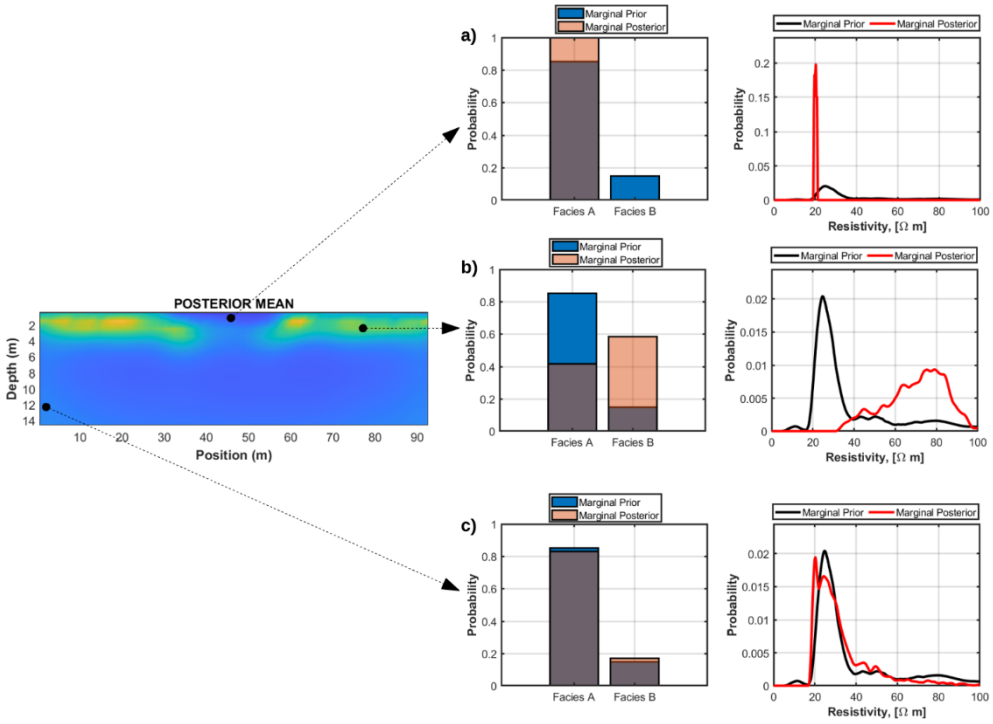

Figure 15: a), b), and c) examples of marginal prior and posterior distributions for the discrete and continuous property at three different spatial locations.

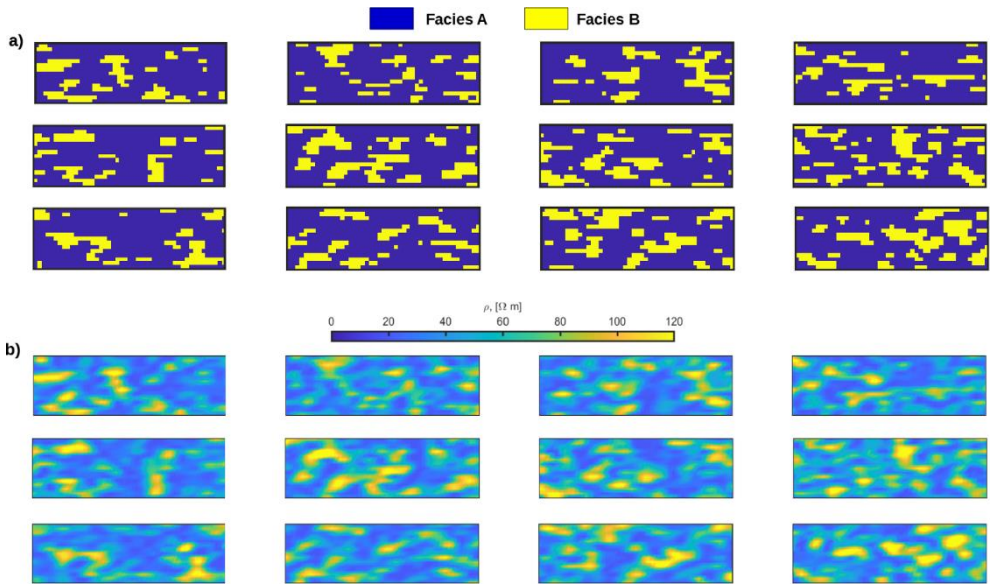

Figure 16: Examples of facies configurations (a) and associated resistivity values (b) extracted from the a-priori distribution and forming 12 out of the 20 current models of the first MCMC iteration.

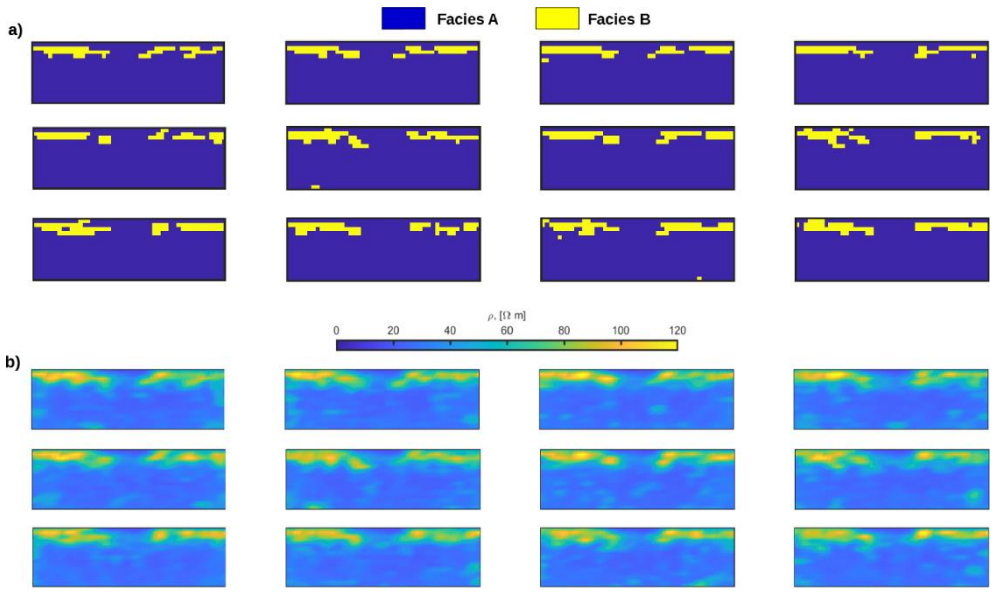

Figure 17: Examples of facies configurations (a) and associated resistivity values (b) forming 12 out of the 20 current models of the last MCMC iteration.

TABLE

	Data error (%)	**Model error (%)**
First iteration	29.9	25.2
Last Iteration	5.8	9.3

Table 1: Data and model errors (in percentage) computed at the first and last MCMC iteration.