



Ensemble-Based Electrical Resistivity Tomography with Data and Model Space Compression

MATTIA ALEARDI,¹  ALESSANDRO VINCIGUERRA,^{1,2} and AZADEH HOJAT^{3,4}

Abstract—Inversion of electrical resistivity tomography (ERT) data is an ill-posed problem that is usually solved through deterministic gradient-based methods. These methods guarantee a fast convergence but hinder accurate assessments of model uncertainties. On the contrary, Markov Chain Monte Carlo (MCMC) algorithms can be employed for accurate uncertainty appraisals, but they remain a formidable computational task due to the many forward model evaluations needed to converge. We present an alternative approach to ERT that not only provides a best-fitting resistivity model but also gives an estimate of the uncertainties affecting the inverse solution. More specifically, the implemented method aims to provide multiple realizations of the resistivity values in the subsurface by iteratively updating an initial ensemble of models based on the difference between the predicted and measured apparent resistivity pseudosections. The initial ensemble is generated using a geostatistical method under the assumption of log-Gaussian distributed resistivity values and a Gaussian variogram model. A finite-element code constitutes the forward operator that maps the resistivity values onto the associated apparent resistivity pseudosection. The optimization procedure is driven by the ensemble smoother with multiple data assimilation, an iterative ensemble-based algorithm that performs a Bayesian updating step at each iteration. The main advantages of the proposed approach are that it can be applied to nonlinear inverse problems, while also providing an ensemble of models from which the uncertainty on the recovered solution can be inferred. The ill-conditioning of the inversion procedure is decreased through a discrete cosine transform reparameterization of both data and model spaces. The implemented method is first validated on synthetic data and then applied to field data. We also compare the proposed method with a deterministic least-square inversion, and with an MCMC algorithm. We show that the ensemble-based inversion estimates resistivity models and associated uncertainties comparable to those yielded by a much more computationally intensive MCMC sampling.

Keywords: Electrical resistivity tomography, Bayesian inversion.

1. Introduction

Electrical resistivity tomography (ERT) is one of the most widely used geophysical methods that can successfully be employed for geotechnical characterization, monitoring earthen dams and embankments, landfill monitoring, groundwater exploration, and mapping of contaminant plumes (e.g., Aleardi et al., 2020; Arosio et al., 2017; Bièvre et al., 2018; Chambers et al., 2006; Dahlin, 2020; Hermans & Paepen, 2020; Hojat et al., 2019a; Karimi-Nasab et al., 2011; Loke et al., 2020; Moradipour et al., 2016; Müller et al., 2010; Pollock & Ciprka, 2012; Supper et al., 2014; Tresoldi et al., 2019; Whiteley et al., 2017). The ERT inversion is nonlinear and ill-posed, thus meaning that the solution to this problem is nonunique due to the limited resolution and noise contamination in the acquired data. The methods used to solve the ERT problem can be classified into two categories: deterministic and probabilistic algorithms. The former are gradient-based approaches that minimize an error function through an iterative linearization procedure (Karoulis et al., 2014; Pidlisecky & Knight, 2008; Zhang et al., 2005). To stabilize the problem, the error function usually includes both a data misfit term and some type of model regularizations, such as smoothness constraints or closeness of the solution to a predefined resistivity model. Undoubtedly, these approaches are the most popular methods to solve the ERT inversion because they guarantee rapid convergence toward a best-fitting model. However, the loss of information produced by the local linearization

¹ Earth Sciences Department, University of Pisa, Via S. Maria 53, 56126 Pisa, Italy. E-mail: mattia.aleardi@unipi.it

² Earth Sciences Department, University of Florence, Via. G. La Pira 4, 50121 Florence, Italy.

³ Department of Mining Engineering, Shahid Bahonar University of Kerman, Jomhouri Blvd., Kerman 76188, Iran.

⁴ Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy.

hinders accurate uncertainty appraisals. Furthermore, these local strategies can often get trapped in local minima of the error function.

On the contrary, probabilistic methods express the inverse solution through the posterior probability density (PPD) function in the model space that fully quantifies the uncertainty on the estimated parameters (Tarantola, 2005). The main advantage of probabilistic over deterministic approaches lies in their ability to provide a set of model realizations that reproduce the measured data. The PPD can be expressed in a closed-form only for linear forward operators and Gaussian assumptions about data and model parameter distributions. For nonlinear problems and/or non-Gaussian assumptions, a complete characterization of the PPD is only possible through sampling, and in these contexts, Markov Chain Monte Carlo (MCMC; Sambridge & Mosegaard, 2002; Sen & Stoffa, 2013) algorithms can be used to numerically estimate the target posterior. These methods provide accurate uncertainty assessments but require a considerable computational effort to converge toward stable PPD estimations, especially in large-dimensional model spaces and for expensive forward model evaluations (Aleardi & Salusti, 2020; Aleardi et al., 2017, 2020; Pradhan & Mukerji, 2020; Sajeja et al., 2014).

To alleviate the computational workload for uncertainty assessment in nonlinear inverse problems, an approximate Bayesian method, the ensemble-based inversion can be applied. This approach is a data assimilation algorithm in which the PPD is represented by an ensemble of model realizations obtained by assimilating the observed data. With an underlying Gaussian assumption on model parameters and noise distribution, this method provides an approximate description of the posterior model, which is usually accurate if the model parameters are approximately Gaussian, the forward operator is not highly nonlinear, and the number of models forming the ensemble is large enough. For example, an initial ensemble with a limited number of models might lead to an underestimation of the posterior uncertainty. The main advantage of this approach is the reduced number of forward modeling runs required to get reasonable PPD estimations with respect to MCMC algorithms. Popular ensemble-based methods are the

ensemble smoother (ES; Evensen, 1994), the ensemble Kalman filter (EnKF; Evensen, 2009), and the ensemble smoother with multiple data assimilation (ES-MDA; Emerick & Reynolds, 2013). Some applications of these approaches to solve geophysical inverse problems can be found in Jin et al. (2008), Tveit et al. (2015), Liu and Grana (2018a, 2018b, 2018c), Thurin et al. (2019), Tveit et al. (2020).

In this work, we propose an ensemble-based inversion of ERT data (EB-ERT) that makes use of the ES-MDA algorithm. Differently from the ES, in ES-MDA the observed data are assimilated multiple times with an inflated data covariance matrix. In this context, each data-assimilation step corresponds to a Bayesian updating step in which the conditional distribution of the model parameters is approximated by estimating the conditional mean and covariance matrix from the ensemble of models. In our application, a finite element (FE) code (Karaoulis et al., 2013) constitutes the forward operator that computes the observed data from the subsurface resistivity values.

To reduce the ill-conditioning and the number of unknown of inversion problems, several compression strategies could be adopted such as wavelet transform (Li & Oldenburg, 2003), singular value decomposition (Liu & Grana, 2018b), Legendre polynomials (Aleardi, 2019), convolutional autoencoder (Liu & Grana, 2018a). Here we reduce both the ill-conditioning of the ERT problem and the number of ensemble members needed for reliable uncertainty through a discrete cosine transform (DCT) reparameterization that is a very common method extensively used for image compression (Britanak et al., 2010). The DCT expands a signal (e.g., expressing the subsurface resistivity model) into a series of cosine functions oscillating at different frequencies. The low-order DCT coefficients express most of the variability of the original signal, and the model compression is simply accomplished by zeroing the numerical coefficients of the basis functions beyond a certain threshold. However, each compression technique must be applied taking in mind that part of the information in the original (unreduced) parameter space could be lost in the reduced space and thus, the model parameterization must always constitute a

compromise between model resolution and model uncertainty (Aleardi, 2020; Aleardi & Salusti, 2021; Dejtrakulwong et al., 2012; Fernández-Martínez et al., 2017; Grana et al., 2019; Lochbühler et al., 2014). Similarly, data parameterization must constitute a compromise between good data resolution (our ability to match the observed data) and accurate uncertainty quantification. It should be also noted that the order of the retained non-zero DCT coefficients in the model space determines the wavelength of the recovered solution along different directions (i.e., vertical, horizontal). Therefore, in our application, the DCT also acts as a regularization operator in the model space that guarantees the preservation of the spatial continuity of the resistivity values in the recovered solution.

The EB-ERT is first assessed on synthetic data and then applied to field data. In both cases, a geo-statistical method generates the initial ensemble of models according to a stationary log-Gaussian prior model and a Gaussian variogram. We also compare the results of the proposed approach with those achieved by an MCMC sampling and with a deterministic least-squares inversion. More specifically, we sample the PPD in the reduced DCT domain by employing the Differential Evolution Markov Chain (DEMC; Vrugt, 2016), an MCMC method that improves over the standard random walk Metropolis because it exploits multiple and interactive chains to sample the parameter space. We will demonstrate that if the number of ensemble members is large enough, the EB-ERT inversion provides reasonable estimations of uncertainties comparable to the MCMC algorithm but with a dramatic reduction of the computational workload. We also assess the influence of the number of ensemble members on the quality of the EB-ERT inversion results. As far as the authors are aware, this is the first time that the EB and MCMC methods are compared in the context of ERT inversion.

2. Methodology

2.1. The Bayesian Approach and the Ensemble-Based Inversion

Gradient-based deterministic inversions are aimed at minimizing a misfit function, which usually is a linear combination of data error and a model regularization term. For Gaussian-distributed noise and model parameters, the error function can be written as follows:

$$E(\mathbf{m}) = \left\| \mathbf{C}_d^{-\frac{1}{2}}(\mathbf{d} - G(\mathbf{m})) \right\|_2^2 + \left\| \mathbf{C}_m^{-\frac{1}{2}}(\mathbf{m} - \mathbf{m}_{prior}) \right\|_2^2, \quad (1)$$

where the vectors \mathbf{m} and \mathbf{d} identify the model parameters and the observed data, respectively; \mathbf{C}_d and \mathbf{C}_m are the data and prior model covariance matrices; \mathbf{m}_{prior} is the prior model vector, and G is the forward modeling operator that maps the model into the corresponding data. The minimum of $E(\mathbf{m})$ can be iteratively approached through a local quadratic approximation of the error function around a starting model (Aster et al., 2018; Menke, 2018).

Differently, a Bayesian inversion is aimed at estimating the full posterior distribution in the model space expressed by:

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})p(\mathbf{m})}{p(\mathbf{d})}, \quad (2)$$

where $p(\mathbf{m}|\mathbf{d})$ is the posterior probability density (PPD), $p(\mathbf{d}|\mathbf{m})$ is the data likelihood function, whereas $p(\mathbf{m})$ and $p(\mathbf{d})$ are the a-priori distributions of model parameters and data, respectively. Under Gaussian assumptions for data and model parameter distributions, and for a linear forward operator \mathbf{G} , we can write (Tarantola, 2005):

$$p(\mathbf{m}) \propto \exp\left(-\frac{1}{2}(\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{prior})\right), \quad (3)$$

$$p(\mathbf{d}|\mathbf{m}) \propto \exp\left(-\frac{1}{2}(\mathbf{d} - \mathbf{Gm})^T \mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{Gm})\right), \quad (4)$$

$$\begin{aligned}
p(\mathbf{m}|\mathbf{d}) &\propto \exp\left\{-\left[\frac{1}{2}(\mathbf{m} - \mathbf{m}_{prior})^T \mathbf{C}_m^{-1}(\mathbf{m} - \mathbf{m}_{prior})\right.\right. \\
&\quad \left.\left. + \frac{1}{2}(\mathbf{d} - \mathbf{Gm})^T \mathbf{C}_d^{-1}(\mathbf{d} - \mathbf{Gm})\right]\right\} \\
&\propto \exp(-E(\mathbf{m}))
\end{aligned} \tag{5}$$

Under the previous assumptions, the maximum a posteriori (MAP) solution of the Bayesian approach coincides with the minimizer of $E(\mathbf{m})$ and can be computed as:

$$\mathbf{m}_{MAP} = \underset{\mathbf{m}}{\operatorname{argmin}} E(\mathbf{m}) = \mathbf{m}_{prior} + \mathbf{K}(\mathbf{d} - \mathbf{Gm}_{prior}), \tag{6}$$

where the matrix \mathbf{K} is the so-called Kalman gain given by:

$$\mathbf{m}_{MAP} = \mathbf{C}_m \mathbf{G}^T (\mathbf{G} \mathbf{C}_m \mathbf{G} + \mathbf{C}_d^{-1}), \tag{7}$$

whereas the posterior covariance matrix can be computed as follows:

$$\mathbf{C}_{m|d} = \mathbf{C}_m - \mathbf{K} \mathbf{G} \mathbf{C}_m. \tag{8}$$

However, for nonlinear inverse problems, the forward operator cannot be expressed into a matrix form and thus, the posterior covariance cannot be analytically computed. Moreover, for large dimensional model spaces, the computation of the Kalman filter becomes computationally unfeasible. To overcome these issues, EB methods replace the analytical model covariance with an empirical estimate of the covariance inferred from the ensemble members. For example, the ES simultaneously assimilates the observed data to globally update an initial ensemble of models generated according to the prior:

$$\mathbf{m}_k^u = \mathbf{m}_k^p + \tilde{\mathbf{K}}(\tilde{\mathbf{d}}_k - \mathbf{d}_k^p), \tag{9}$$

with $k = 1, \dots, N$, where N represents the number of models in the ensemble and $\tilde{\mathbf{d}}_k$ is a random perturbation of the observed data according to the Gaussian distribution $\mathcal{N}(\mathbf{d}, \mathbf{C}_d)$. The subscripts u and p represent the updated (current iteration) and prior (previous iteration) variables, respectively; \mathbf{d}_k^p is the data associated with the k th prior model \mathbf{m}_k^p . Finally, $\tilde{\mathbf{K}}$ is the Kalman gain that can be estimated from the ensemble of models according to:

$$\tilde{\mathbf{K}} = \mathbf{C}_{md}^p (\mathbf{C}_{dd}^p + \mathbf{C}_d)^{-1}, \tag{10}$$

where \mathbf{C}_{dd}^p denotes the covariance matrix of the predicted data; \mathbf{C}_{md}^p is the cross-covariance between the model vector \mathbf{m}^p and the vector of the associated data \mathbf{d}^p . It can be demonstrated that the EnKF and ES are equivalent to a single Gauss–Newton iteration. This means that an iterative updating of the prior models is equivalent to performing multiple smaller corrections to the ensemble. The ES usually requires many iterations to achieve good data predictions. For this reason, the ES-MDA was proposed in which the data are assimilated multiple times with an inflated data covariance. In this context, Eq. (9) can be interpreted as a Bayesian updating step under the assumption of Gaussian distributed model parameters. Note that this equation can be also applied to nonlinear problems because the covariance matrices \mathbf{C}_{dd}^p and \mathbf{C}_{md}^p are approximately inferred from the ensemble models and the associated data. The ES-MDA improves over the ES because it reduces the model variations at each iteration by employing a damping parameter α that scales the data covariance matrix. The steps forming the ES-MDA can be described as follows:

1. Define the total number of iterations Q , the number of models in the ensemble N , and the inflation coefficient α for each iteration with $\sum_{i=1}^Q \frac{1}{\alpha_i} = 1$;
2. Generate the prior ensemble of models according to the Gaussian prior;
3. For each i th iteration
 - (a) Run the forward models to compute the data associated with each ensemble member $\{\mathbf{d}^p\}_{1, \dots, N}$;
 - (b) Perturb the observed data for each k th ensemble member according to: $\tilde{\mathbf{d}}_k = \mathbf{d} + \sqrt{\alpha_i} \mathbf{C}_d^{-1/2} \mathbf{n}$, with $\mathbf{n} = \mathcal{N}(0, \mathbf{I})$, where \mathbf{I} is the identity matrix, whereas $\alpha_i \mathbf{C}_d$ denotes the inflated covariance matrix;
 - (c) Update the ensemble using Eqs. (9) and (10) in which \mathbf{C}_d is replaced by $\alpha_i \mathbf{C}_d$.

The algorithm updates each model to minimize the L2 norm difference between the observed data and the data computed on the ensemble members. All

the models that attain the final misfit values can be considered a sample from the equivalence region of solution in the model space (Aster et al., 2018). In other terms, such models equally reproduce the observed data and for this reason, represent possible subsurface scenarios in agreement with the acquired geophysical data and with the prior assumptions. From this ensemble of models, we can numerically derive an estimate of the posterior uncertainty.

the one-dimensional definition because it is simply a separable product (equivalently, a composition) of DCTs along each dimension. For example, if we assume a 2-D resistivity model $\rho(x, y)$ in which $x = [0, 1, \dots, M_x - 1]$ and $y = [0, 1, \dots, M_y - 1]$ represent the horizontal and vertical coordinates, respectively, the associated 2-D transform is defined as follows:

$$\begin{cases} \mathbf{R}(k_x, k_y) = \frac{1}{\sqrt{M_x}} \frac{1}{\sqrt{M_y}} \sum_{x=0}^{M_x-1} \sum_{y=0}^{M_y-1} \rho(x, y), & \text{if } k_x = k_y = 0 \\ \mathbf{R}(k_x, k_y) = \sqrt{\frac{2}{M_x}} \sqrt{\frac{2}{M_y}} \sum_{x=0}^{M_x-1} \sum_{y=0}^{M_y-1} \rho(x, y) \cos\left(\frac{(2x+1)\pi k_x}{2M_x}\right) \cos\left(\frac{(2y+1)\pi k_y}{2M_y}\right), & \text{if } k_x, k_y \neq 0 \end{cases}, \quad (11)$$

The number of ensemble members should be large enough to get an accurate estimate of the \mathbf{C}_{dd}^p and \mathbf{C}_{md}^p matrices but small enough not to make the forward evaluations computationally impractical. The optimal number of members also depends on the model space dimension: large model spaces require a considerable number of ensemble members to achieve reliable uncertainties estimations. Indeed, it is known that too small ensembles can result in an underestimation of the model uncertainties. To overcome these issues and to also mitigate the ill-conditioning of the problem we compress both data and model spaces using the DCT.

2.2. Discrete Cosine Transform

Several variants of DCT exist with slightly modified definitions, but in this work, we use the so-called DCT-2 formulation that is the most common. Hereafter we simply refer to the DCT-2 transformation as the DCT. We employ the DCT parameterization because it exhibits superior compression power over other compression methods (Lochbühler et al., 2014).

The DCT can be applied to multidimensional signals (i.e., 2-D matrices) and such multi-dimensional DCT transform follows straightforwardly from

where $\mathbf{R}(k_x, k_y)$ represents the k_x -th and k_y -th DCT coefficient. The values within the matrix \mathbf{R} represent the unknowns to be estimated in a DCT-reparameterized inverse problem. Equation (11) can be compactly rearranged in matrix form:

$$\mathbf{R} = \mathbf{B}_y \rho \mathbf{B}_x^T, \quad (12)$$

where \mathbf{B}_x and \mathbf{B}_y are the matrices with dimensions $M_x \times M_x$ and $M_y \times M_y$, respectively that contain the DCT basis functions, whereas the $M_y \times M_x$ matrix \mathbf{R} expresses the DCT coefficients. This approach concentrates most of the information of the original signal into the low-order DCT-coefficients, and hence, an approximation of the subsurface resistivity model can be obtained as follows:

$$\tilde{\rho} = \left(\mathbf{B}_y^q\right)^T \mathbf{R}_{qp} \mathbf{B}_x^p, \quad (13)$$

where $\tilde{\rho}$ is the approximated $[M_y \times M_x]$ resistivity model, \mathbf{B}_y^q is a $[q \times M_y]$ matrix containing only the first q rows of \mathbf{B}_y ; \mathbf{B}_x^p is a $[p \times M_x]$ matrix containing only the first p rows of \mathbf{B}_x , whereas the matrix \mathbf{R}_{qp} represents the first q rows and p columns of \mathbf{R} . In other words, the scalar q and p represent the retained number of basis functions along the y and x directions used to derive the approximated model. Therefore, the DCT transformation allows a reduction of the $(M_y \times M_x)$ -D full resistivity model space to a $(q \times p)$ -

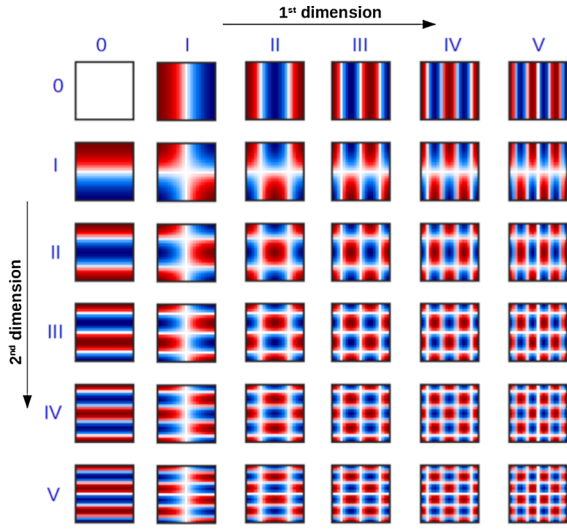


Figure 1

2-D DCT basis functions of different orders. Blue and red code low and high numerical values, respectively

D DCT-compressed parameter space with $p < M_x$ and $q < M_y$. In this context the $p \times q$ non-zero numerical coefficients of the \mathbf{R}_{qp} matrix become the unknowns to be estimated after a DCT compression of the model space. Estimating the retained DCT-coefficients reduces the parameter space dimensionality and can significantly improve the computational efficiency of the inversion procedure. Figure 1 shows

some DCT basis functions of different orders in a 2-D space and illustrates that the variability of the solution along each dimension is directly determined by the orders of the retained DCT coefficients. Note that moving from the full data and model spaces to the DCT-compressed domains reduces both the numbers of ensemble members needed for reliable uncertainty estimation and the dimensions of the \mathbf{C}_{dd}^p and \mathbf{C}_{md}^p matrices, thereby reducing the computational effort of the entire inversion procedure.

In this work, we apply the 1D DCT to compress the measured apparent resistivity values, and the 2D DCT to compress the subsurface resistivity model. This means that the model parameter vector \mathbf{m} includes the retained number of coefficients in the model space, while the observed data vector \mathbf{d} includes the retained coefficients in the data space. Figure 2 shows a schematic representation of the projection of the full model and data spaces onto the compressed DCT domain and the so derived model and data vectors for the EB-ERT inversion. Note that because of its trapezoidal shape, the apparent resistivity pseudosection cannot be expressed with a 2D matrix and thus, it has been flattened to a 1D vector before the DCT projection. After generating the initial ensemble according to the prior assumptions and computing the associated predicted data, all the

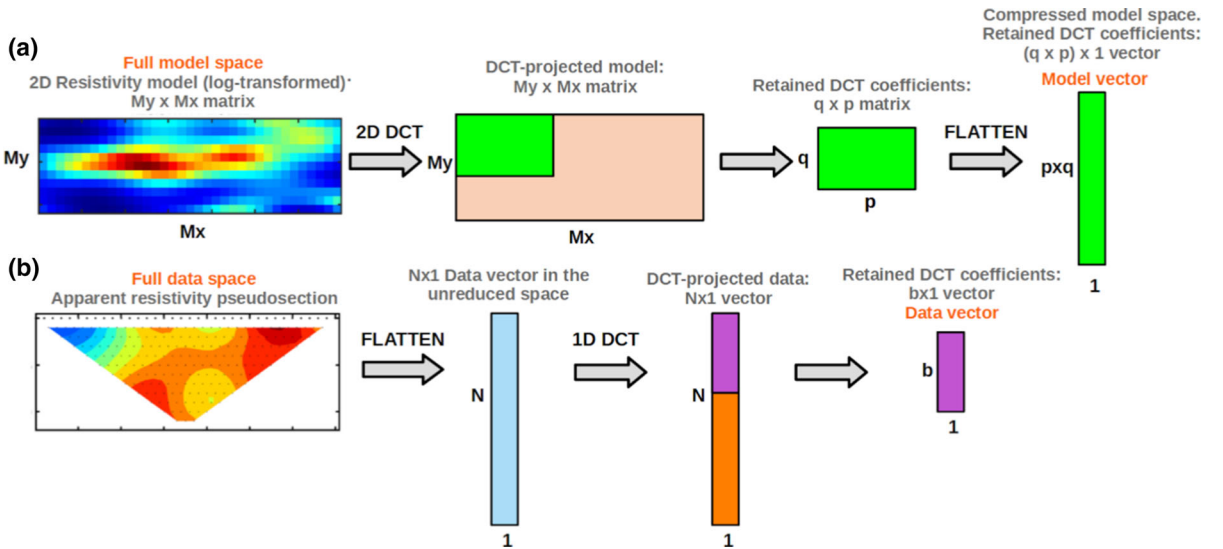


Figure 2

Derivation of the model (a) and data (b) vectors in the DCT space from the resistivity model and the associated pseudosection

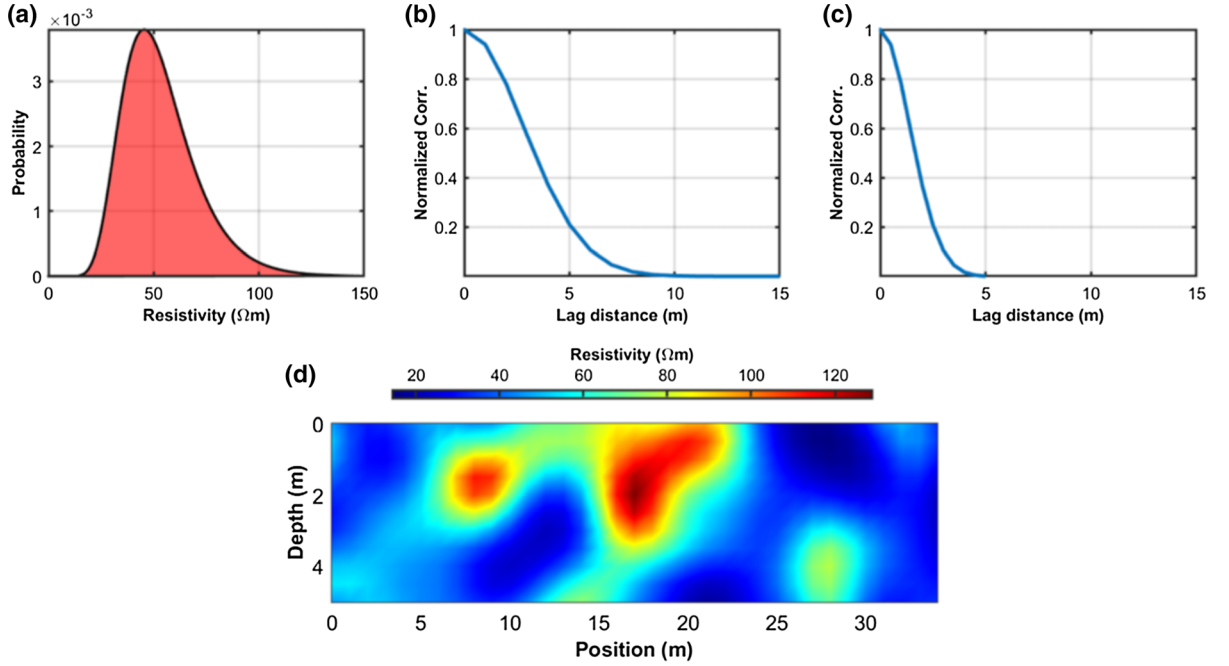


Figure 3

a Marginal log-Gaussian prior distribution for the resistivity values. **b**, **c** represent the spatial correlation functions associated with the assumed 2-D variogram model along the lateral and vertical directions, respectively. **d** The true model for the synthetic inversion

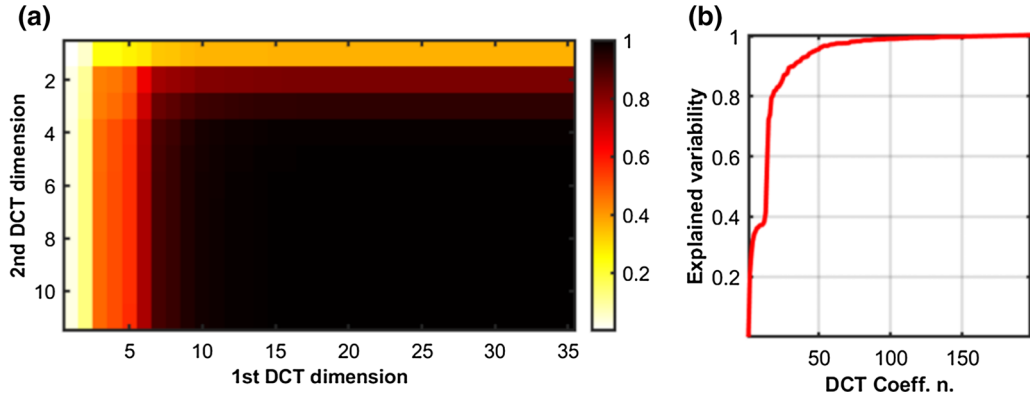


Figure 4

a Examples of explained model variability as the number of DCT coefficients along the 1st and 2nd DCT dimension increases. We illustrate the explained variability for a resistivity model drawn from the prior distribution. The numerical value with coordinate (x, y) indicates the explained variability if the first x , and y DCT coefficients along the 1st and 2nd DCT dimensions, respectively, are used for compressing the resistivity model. It emerges that 10 DCT coefficients along the 1st dimension and 4 along the 2nd dimension explain almost the 100% of the variability of the uncompressed resistivity model. **b** Explained variability as the number of DCT coefficients increases for the data associated with the model shown in **a**. In this case, 80 DCT coefficients express almost the total variability of the uncompressed dataset

models and data are projected onto the DCT space. We will consider log-Gaussian distributed resistivity values and for this reason, the DCT coefficients to be

estimated refer to the natural logarithm of the resistivity. The inversion runs in the reduced domain and after the selected number of iteration has been

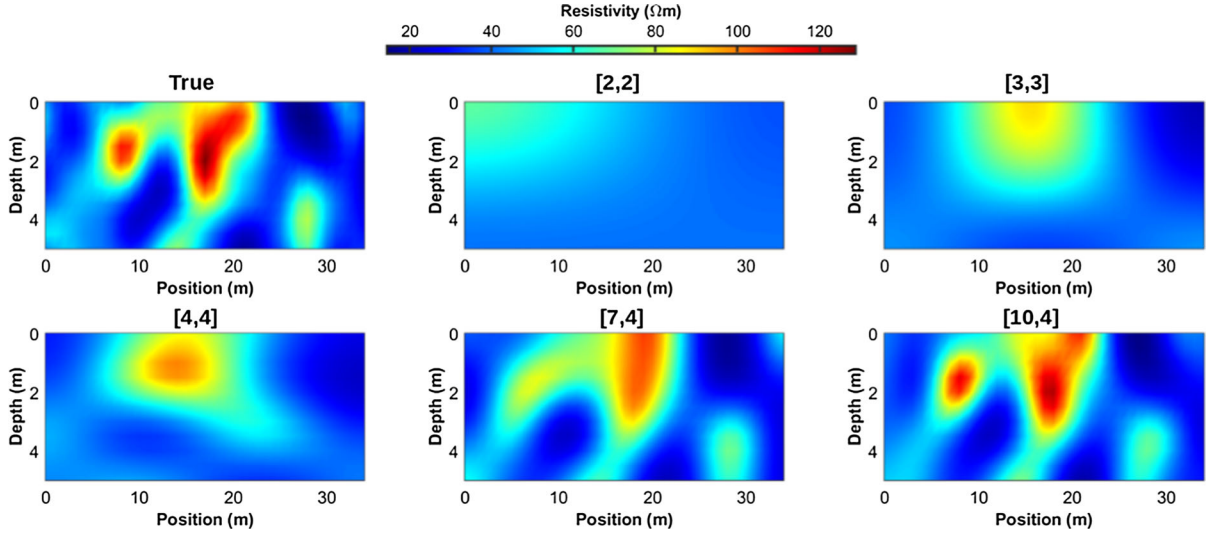


Figure 5

A comparison between the true resistivity model drawn from the prior distribution and some approximations obtained with different numbers of DCT coefficients. The parentheses express the number of coefficients retained along the first and second DCT dimensions. For example [7, 4] indicates that we consider 7 and 4 coefficients along the first and second DCT directions, respectively

reached, the ensemble of DCT models collected at the last iteration are projected back onto the resistivity space to compute the mean model and the associated uncertainty in the original, uncompressed domain.

3. Synthetic Inversion

In the following synthetic example, the study area is 35 m long and 5.5 m deep and is discretized with rectangular cells with spatial dimensions of 1 m and 0.5 m along the horizontal and vertical directions, respectively. We define a stationary log-Gaussian prior model and a spatial Gaussian variogram that can be expressed by the following correlation function:

$$\tau_x = \exp\left(-\frac{h_x^2}{\alpha_x^2}\right), \quad (14)$$

where τ_x indicates the correlation along the horizontal x -direction, h_x is the spatial distance of the autocorrelation function along the same direction, and α_x is the parameter that defines the spatial dependency along the x -axis. Note that the scalar value α_x directly influences the variogram range along the x -direction. In this example we consider α_x and α_y equal to 4 and 2 m, respectively. The marginal log-Gaussian prior

Table 1

Quantitative summary of the EB inversion results

N. models	CC model	RMSE model	CC data	RMSE data	Cov. ratio	Time to converge (s)
100	0.75	16.43	0.98	4.55	0.65	90
200	0.78	15.30	0.98	4.57	0.71	150
500	0.80	14.61	0.98	4.59	0.83	360
1000	0.80	14.65	0.98	4.60	0.85	720
2000	0.80	14.63	0.98	4.58	0.86	1440

From left to right: the number of models forming the ensemble (N. models). Correlation coefficient and RMSE between the true model and the estimated mean model (CC model, and RMSE model, respectively). Correlation coefficient and RMSE between the observed data and the data generated by the mean model (CC data, and RMSE data, respectively). Coverage ratio for the 80% confidence interval. Time to converge. Note that the comparison between the true and predicted model and between the predicted and observed data is performed in the linear domain and not in the logarithmic domain

together with the lateral correlation functions are represented in Fig. 3. The true model is simply generated as a random realization extracted from the prior assumptions. This means that we put ourselves in the most favorable case in which the prior assumptions correctly capture the statistical

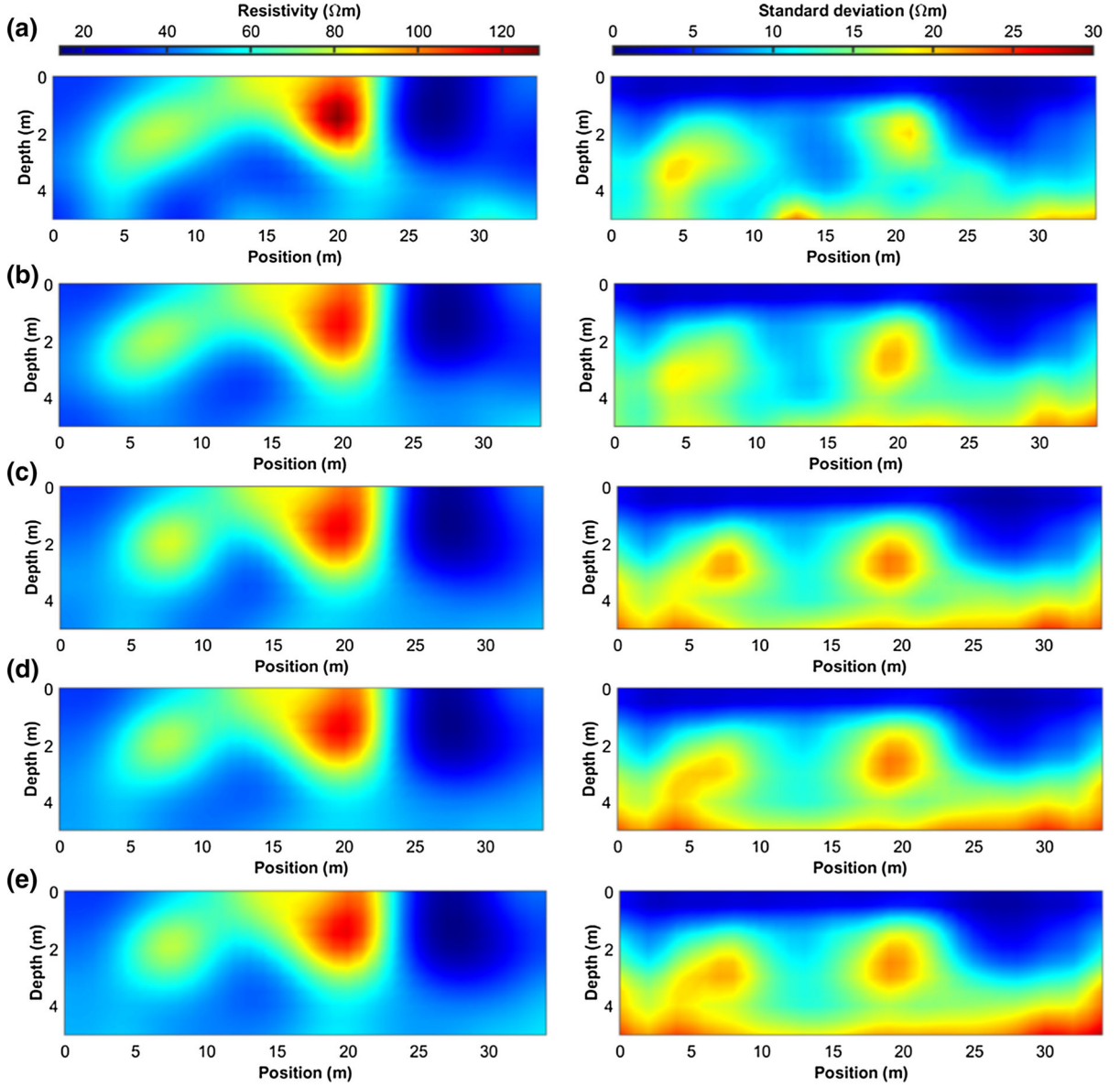


Figure 6

Estimated mean models (left) and associated standard deviation maps (right) for the different tests. **a** 100 models in the ensemble. **b** 200 models in the ensemble. **c** 500 models in the ensemble. **d** 1000 models in the ensemble. **e** 2000 models in the ensemble

properties and spatial variability patterns of the sub-surface resistivity values. The true model includes two high-resistivity anomalies embedded in a low resistivity medium: one wider anomaly is located at the center of the model, while the other is located around the horizontal coordinate of 7.5 m. At the right edge of the model (between the horizontal

coordinates of 25 and 30 m) we have also simulated a shallow and low-resistivity body.

We simulate a Wenner acquisition layout with 36 electrodes evenly spaced of 1 m and an injected current of 1 A. The maximum a value is 11. This configuration results in $11 \times 35 = 385$ model parameters to be estimated from 198 data points. The previously mentioned FE code has been used to

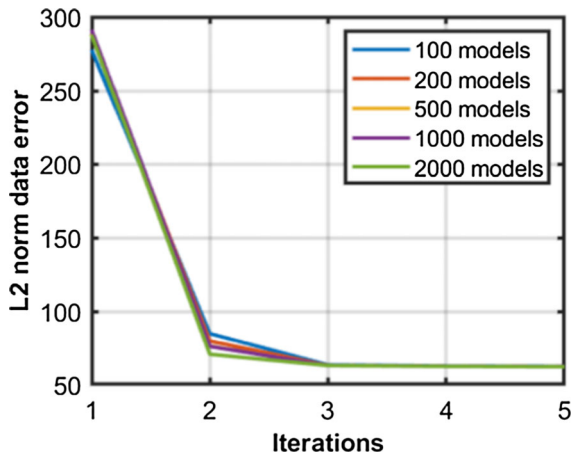


Figure 7

Evolution of the L2 norm data misfit for the different tests

compute the noise-free observed data that have been contaminated with uncorrelated Gaussian noise with a standard deviation equal to the 20% of the total standard deviation of the noise-free dataset. This results in a signal-to-noise ratio of 20 dB.

After defining the prior we must determine the optimal number of DCT coefficients needed to approximate the subsurface resistivity model. To this end, we quantify how the explained variability of the resistivity model changes as the number of DCT basis functions increases. The selection of the optimal number of DCT coefficients is a very delicate step that must guarantee uncertainty estimations, model resolution, and data fitting comparable to those achieved by an inversion running in the full, uncompressed space. A detailed discussion on how the model and data compressions affect the inversion results is far beyond the scope of this work and we refer the interested reader to Grana et al. (2019) for more theoretical insights. Figure 4a shows the explained variability for a model drawn from the prior as the number of retained DCT coefficients increases. We note that only 40 coefficients (i.e., 10 and 4 coefficients along the first and second DCT directions, respectively) explain almost the total variability of the uncompressed resistivity model. This means that the DCT allows for a reduction of the

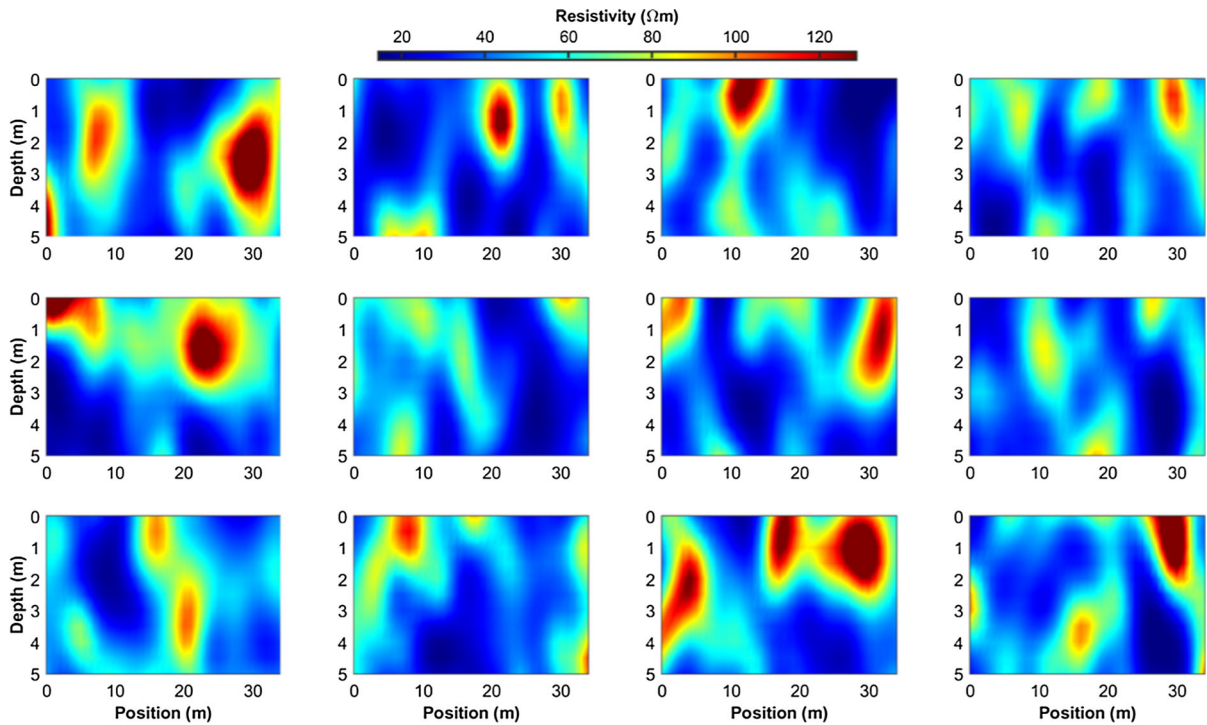


Figure 8

Some examples of resistivity sections drawn from the prior distribution and associated with the test employing 500 models in the ensemble

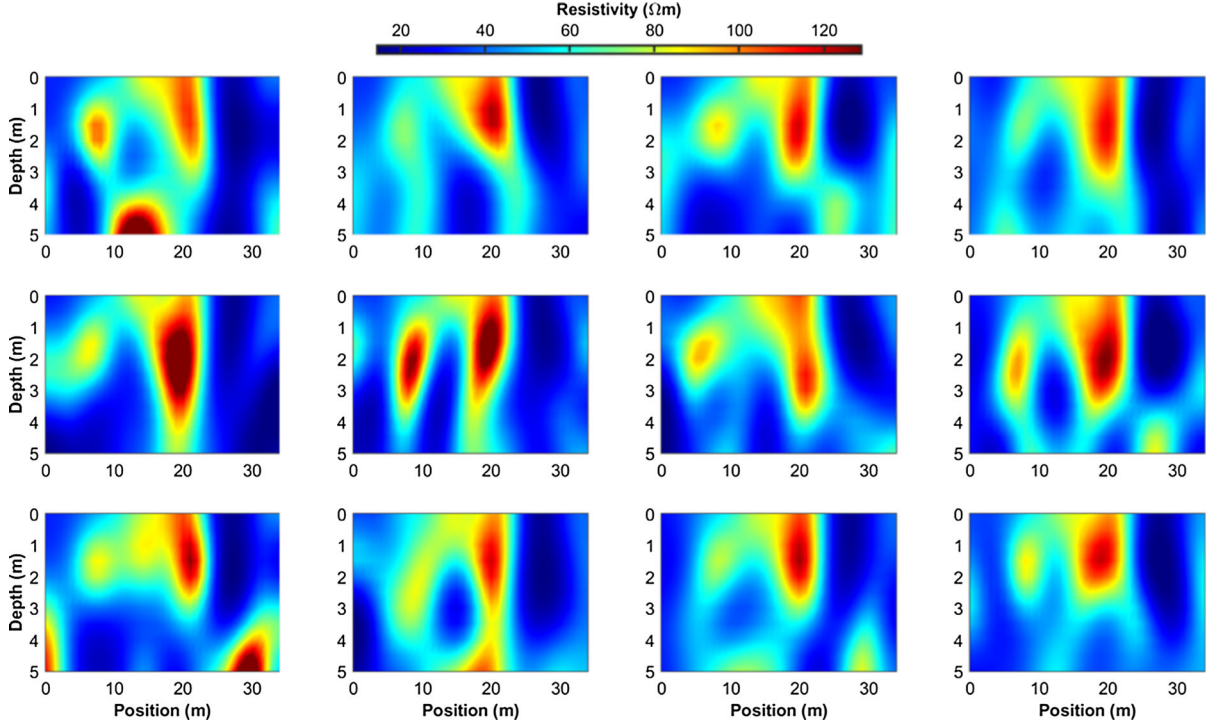


Figure 9

Some examples of models forming the ensemble at the third iteration for the test with 500 models

385D model space to $4 \times 10 = 40$ D domain. A similar analysis has been carried out on some datasets derived from prior realizations. An example is shown in Fig. 4b where it emerges that 80 coefficients successfully compress almost the total variability of the uncompressed dataset. Therefore, in this case, the full 385D model space has been reduced to a 40D domain, while the full data space has been reduced to an 80D domain. As an example, Fig. 5 compares the true resistivity model with the approximated models derived for different numbers of retained DCT coefficients. If only two coefficients are considered along the two DCT dimensions, the actual resistivity values are not recovered. Four coefficients along the two DCT dimensions provide just a very smoothed version of the original model, while only 10 and 4 coefficients guarantee an almost complete reproduction of the lateral and vertical variations of the actual resistivity values. Therefore, the reduced spatial resolution that we will observe in the synthetic results is only related to the acquisition layout we simulate and not to the number of DCT coefficients employed.

In the following synthetic experiments, we perform 5 tests employing different numbers of ensemble members with the aim to assess how the dimension of the ensemble affects the recovered model and the associated uncertainty. In all cases, the ensemble is evolved for 5 iterations. The simple kriging algorithm is used to generate the initial ensemble according to the assumed variogram model and the prior Gaussian distribution in the logarithm domain. We implement Matlab codes in which the generation of the initial ensemble and the forward evaluations are distributed across different cores to reduce the computational workload. The inversions are run on two deca-core intel E5-2630 at 2.2 GHz (128 Gb RAM). We measure the accuracy of the results using the linear correlation coefficient and root-mean-square error (RMSE) between the true model and the mean estimated model, and between the observed data and the data predicted on the mean model. We also assess the estimated posterior quantification by computing the coverage ratio for the 80% confidence interval (Table 1). We also evaluate

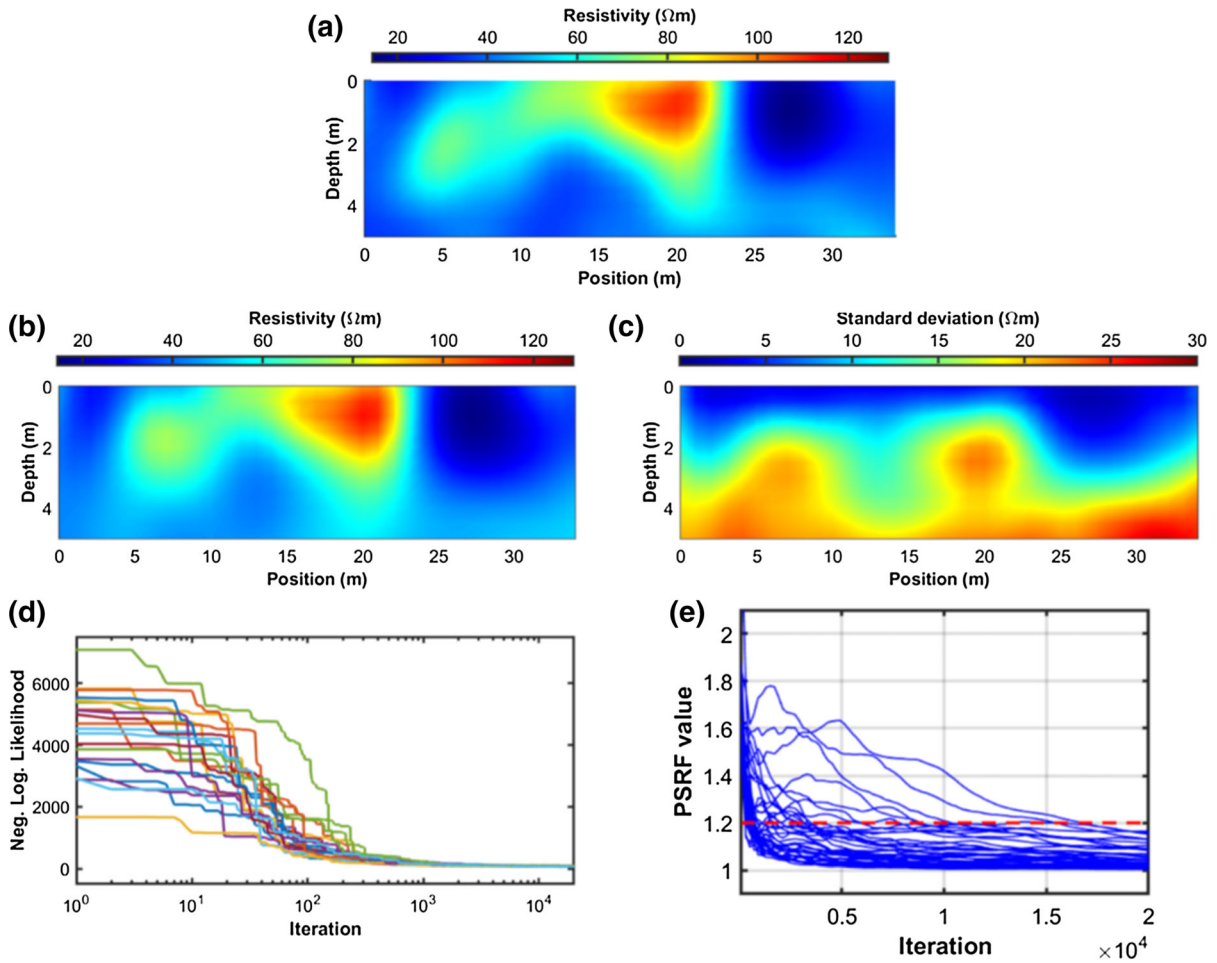


Figure 10

a Predicted model provided by the deterministic inversion. **b** Posterior mean model estimated by the MCMC inversion. **c** posterior standard deviation provided by the MCMC algorithm. **d** Evolution of the negative log-likelihood for the 20 chains. **e** Evolution of the potential scale reduction factor for the 40 model parameters

the computing time needed to achieve a stable data misfit value.

The estimated mean ensemble models and the associated standard deviations are represented in Fig. 6. We observe that in all cases we get similar estimates for the mean model: the wide anomaly in the central part of the investigated area is well recovered, as well as the low resistivity body located in the shallowest part and between the horizontal coordinates of 25–30 m. The actual resistivity value associated with the deeper and smaller anomaly is always underestimated. In all cases, as expected, the uncertainty increases moving from the shallow to the

deep and lateral parts of the model. The uncertainty also increases as the resistivity value increases. However, for 100 and 200 models in the ensemble, we obtain an underestimation of the uncertainties compared with the other tests. This underestimation increases as the data illumination decreases and for this reason, it is more significant at the bottom and at the lateral edges of the study area. Differently, the estimated uncertainty is similar for the tests with 500, 1000, and 2000 models. Table 1 indicates that all the tests provide a mean model that closely reproduces the observed data. The correlation coefficients and the RMSE in the model space indicate that the

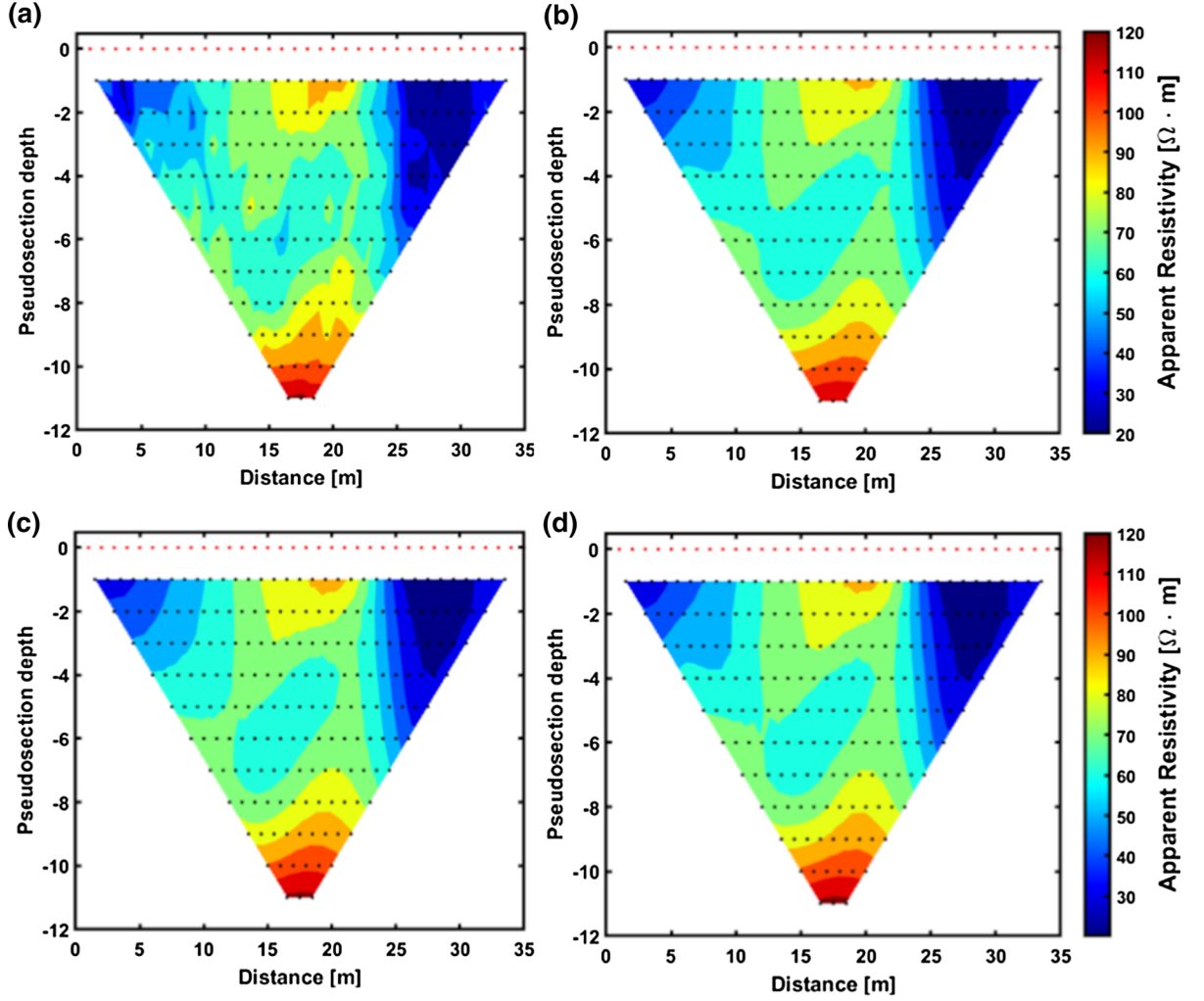


Figure 11

a Observed pseudosection. **b** Predicted data computed on the model estimated by the deterministic inversion (b), on the mean model estimated by the EB (c), and on the mean model provided by the MCMC inversion (d)

accuracy of the results does not significantly change for ensembles with 500 or more models. Similarly, the coverage ratio shows that the precision of the results does not change in the last three tests with 500, 1000, and 2000 models. The computational cost per iteration is 29 s when 100 models are considered for the ensemble and it is increased up to 480 s when we consider 2000 models. Figure 7 represents the evolution of the L2 norm data misfit values for different tests and over 5 iterations. In all cases, the same final error value is reached in just 3 iterations. This means that the total computing time goes from

90 s for an ensemble of 100 models up to 1440 s (24 min) for 2000 ensemble members. However, we also note that larger ensembles are associated with a faster convergence rate in the very first iterations because they more accurately approximate the C_{dd}^p and C_{md}^p matrices. These results indicate that for this test an ensemble of 500 models constitutes the best compromise between the accuracy and precision of the results and the computational cost of the EB-ERT inversion (i.e., 6 min to attain convergence). Figures 8 and 9 compare some resistivity sections associated with the test employing 500 ensemble

Table 2

Quantitative summary of the deterministic and MCMC inversion results

	CC model	RMSE model	CC data	RMSE data	Cov. ratio	Time to converge
Deterministic	0.80	14.62	0.98	4.53	–	2 min
MCMC	0.80	14.65	0.98	4.58	0.89	74 h

From left to right: Correlation coefficient and RMSE between the true model and the estimated mean model (CC model, and RMSE model, respectively). Correlation coefficient and RMSE between the observed data and the data generated on the mean model (CC data, and RMSE data, respectively). Coverage ratio for the 80% confidence interval. Total computing time to converge toward a stable PPD

members: Fig. 8 represents some models sampled from the prior, whereas Fig. 9 shows some resistivity sections forming the ensemble at the third iteration. In the latter case, we note that the resistivity anomalies are univocally recovered by all the models.

We now compare the outcomes of the EB-ERT inversion with those provided by an MCMC sampling running in the compressed space and with a standard, least-squares deterministic inversion running in the full data and model spaces. We implement a DEMC algorithm running for 20,000 iterations and employing 20 chains. The prior is the same used for the EB inversion. The deterministic inversion provides a

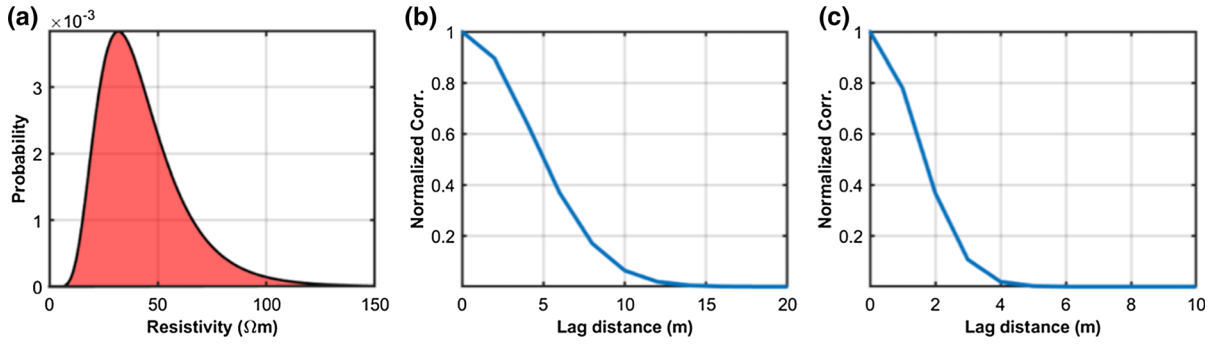


Figure 12

a Marginal log-Gaussian prior distribution for the resistivity values in the real data application. **b, c** The spatial correlation functions associated with the assumed 2-D variogram model along the lateral and vertical directions, respectively

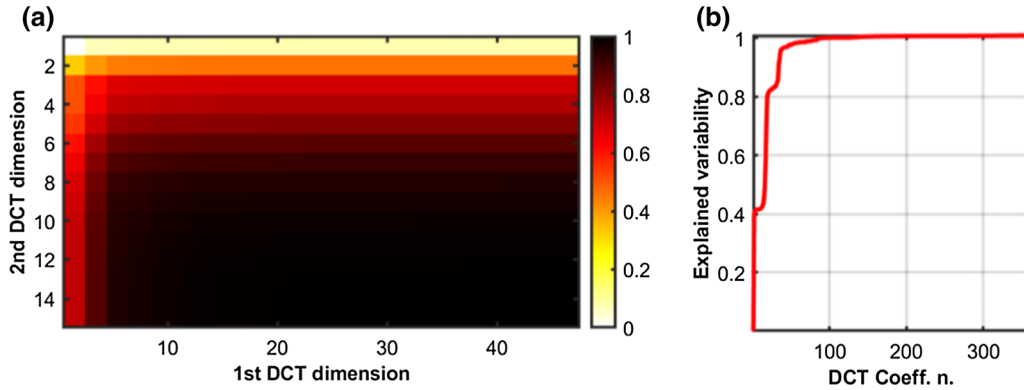


Figure 13

a Examples of explained model variability as the number of DCT coefficients along the 1st and 2nd DCT dimension increases. We illustrate the explained variability for a resistivity model drawn from the prior distribution. The numerical value with coordinate (x, y) indicates the explained variability if the first x , and y DCT coefficients along the 1st and 2nd DCT dimensions, respectively, are used for compressing the resistivity model. It emerges that 15 DCT coefficients along the 1st dimension and 10 along the 2nd dimension explain almost the 100% of the variability of the uncompressed resistivity models. **b** Explained variability as the number of DCT coefficients increases for the data associated with the model shown in **a**. In this case, 150 DCT coefficients express the total variability of the uncompressed data

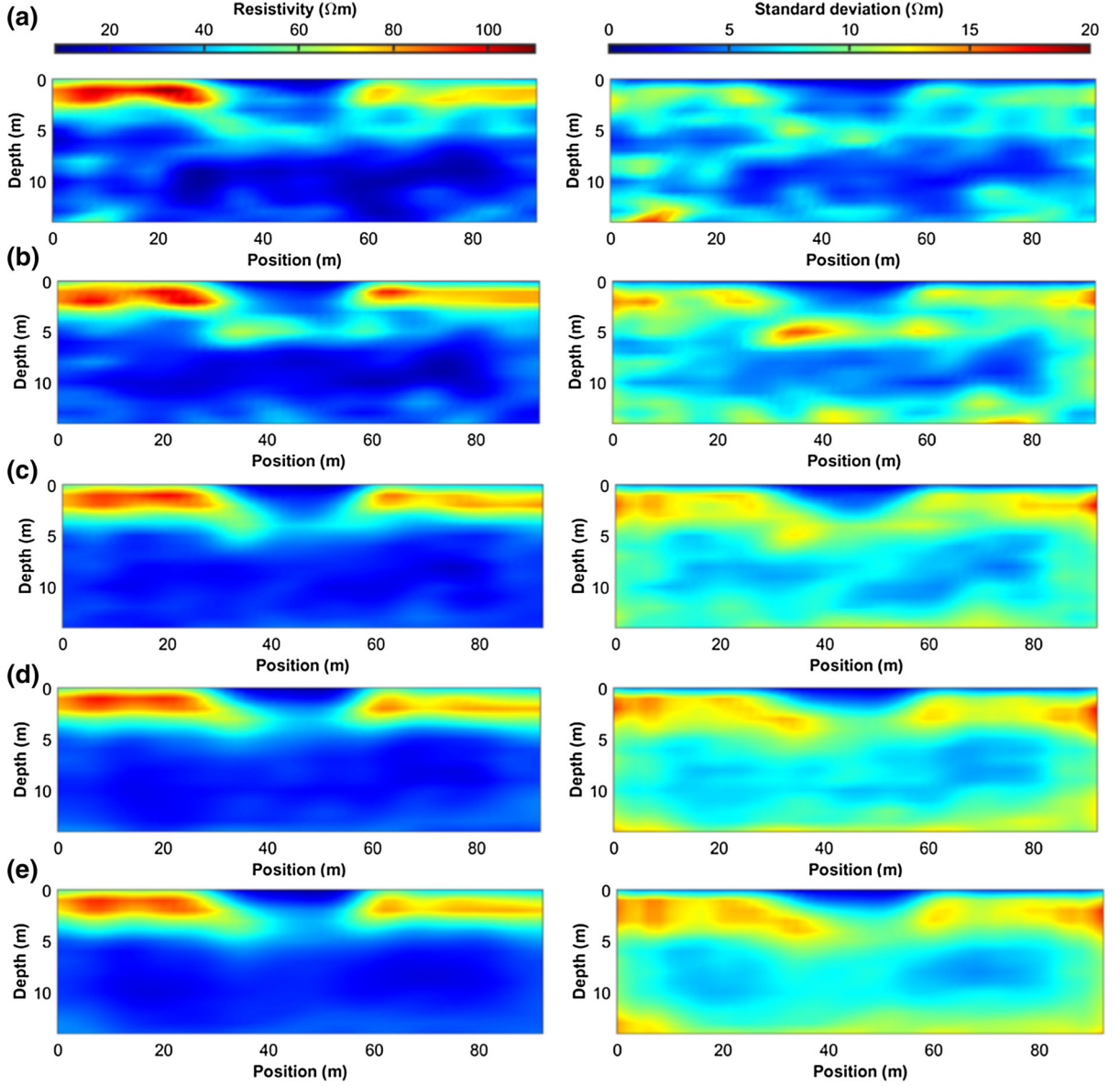


Figure 14

Estimated mean models (left) and associated standard deviation maps (right) for different tests. **a** 100 models in the ensemble. **b** 200 models in the ensemble. **c** 500 models in the ensemble. **d** 1000 models in the ensemble. **e** 2000 models in the ensemble

final prediction very similar to the posterior model estimated by the MCMC algorithm (Fig. 10a, b). Both these predicted models and the standard deviation estimated by the MCMC inversion (Fig. 10a, b) are similar to those provided by the EB-ERT approach with more than 500 members forming the ensemble, although we observe a slight increase of the posterior uncertainties where the illumination is

poor (e.g., at the bottom of the model; Fig. 10c) with respect to the EB results (Fig. 6c). The evolution of the negative log-likelihood illustrates that the stationary regime for the MCMC algorithm is attained in 3000 iterations, approximately (Fig. 10d). To quantitatively assess the convergence of the sampling we use the potential scale reduction factor (PSRF), a popular MCMC convergence diagnostic tool

Table 3

Computing time to converge toward a stable data misfit value for the field inversion tests with different number of ensemble members

N. models	Time to converge (s)
100	111
200	213
500	519
1000	960
2000	2040

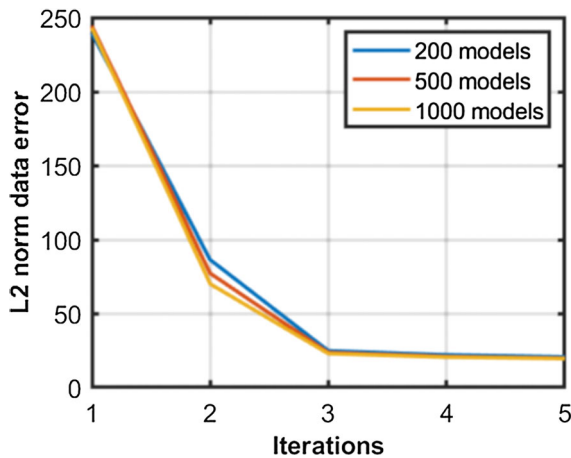


Figure 15

Evolution of the L2 norm data misfit for the tests with 200, 500, and 1000 models in the ensemble

proposed by Brooks and Gelman (1998) to which we refer the reader for its formal definition. PSRF compares within-chain variances to the variance computed from all mixed chains for a given parameter. In practice, one can consider that the convergence to a stable posterior model has been achieved if the PSRF is lower than 1.2. In Fig. 10e we observe that more than 17,000 iterations are needed to achieve convergence for all the 40 model parameters. This translates into a total computational cost for the MCMC inversion of more than three days, approximately, using a parallel Matlab code running on the previously described hardware configuration. Figure 11 shows the comparison between the observed data and the data generated on the deterministic results and on the mean models estimated by the EB and MCMC inversion. In this synthetic example, all the approaches provide final

predictions that closely reproduce the observed pseudosection.

Table 2 summarizes the results achieved by the deterministic and MCMC inversion. In both cases, we note that the RMSE and correlation coefficients are congruent with those obtained by the EB approach. The coverage ratio can not be determined for the deterministic inversion because this method does not provide model uncertainties. The slightly higher coverage ratio provided by MCMC demonstrates that this method gives more accurate uncertainty estimations than the EB approach but at the expense of a dramatic increase of the computational effort of more than 3 orders of magnitude. Table 2 also shows that the computational cost of the deterministic inversion is negligible in comparison to that of the MCMC sampling. From this synthetic test emerges that the EB method tends to slightly underestimate the uncertainties in the least illuminated parts of the subsurface. However, from a practical point of view, we deem that the model uncertainties estimated by the EB inversion are reasonable and comparable to those yielded by the Monte Carlo sampling.

4. Field Data Inversion

We now discuss the application to 2D field data acquired by a permanent system installed along a critical part of Parma river embankment in Colorno, Italy. 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 and 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. We employ 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. The data have been corrected for the effect of the soil overlaying electrodes (Hojat et al., 2020).

Based on the available information about the study area we define a log-Gaussian prior resistivity model and a spatial variability pattern described by a

Gaussian variogram with α_y and α_x values equal to 2 and 6 m, respectively (Fig. 12). From previous studies on this investigated site, we expect a mainly low-resistivity clay body that, at the two lateral parts of the section and around 2–3 m depth, hosts a more permeable layer constituted by sand and gravel and characterized by higher resistivity values.

To determine the optimal number of DCT coefficients in the model and data spaces we again generate some prior realizations and compute the associated data, to evaluate how the explained variability changes as the number of retained coefficients increases. From Fig. 13a we observe that 15 and 10 coefficients along the first and second DCT dimensions, respectively, explain almost the total variability of the uncompressed model, whereas Fig. 13b illustrates that 150 coefficients recover the total variability in the data space. Therefore, in this case, the DCT transformation allows reducing the full 705D model and 360D data spaces into compressed 150D domains.

Again, we run different tests that employ different numbers of ensemble members and we compare the different mean models and PPD estimations to select the best compromise between the computational cost and the accuracy and precision of the final results. The FE code previously used in the synthetic inversion again constitutes the forward modeling engine for the inversion process. Similar to the synthetic example, we run 5 iterations for each test.

Figure 14 shows the estimated mean models and the standard deviations for the different tests. In all cases, and as expected from previous information about the study site, we predict the two lateral high resistivity bodies around 2 m depth (associated with sand/gravel) hosted in a low resistivity medium (clay). We also note that the estimated mean model and the associated uncertainties stabilize for the tests employing 1000, and 2000 members, while the tests with only 100, 200, and 500 models provide estimated mean with some artifacts and underestimated uncertainties. However, as expected, all the tests yield uncertainty estimations in which the shallowest

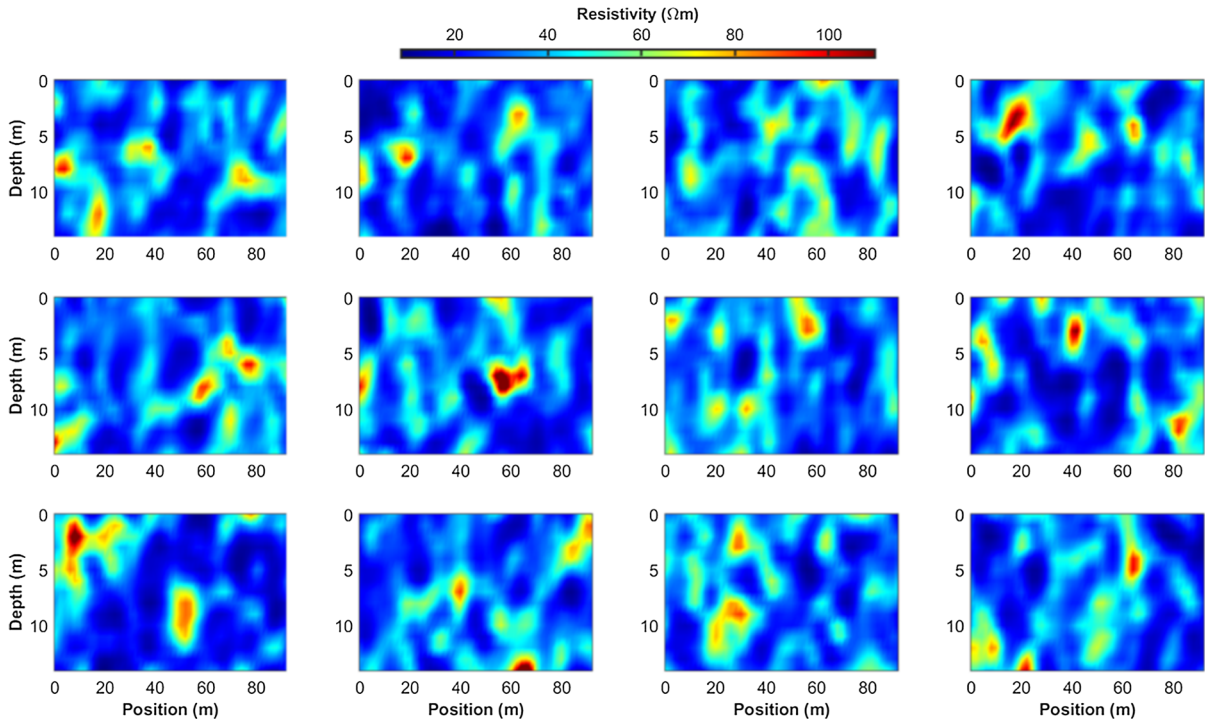


Figure 16

Some examples of resistivity sections drawn from the prior distribution and associated with the test employing 1000 models in the ensemble

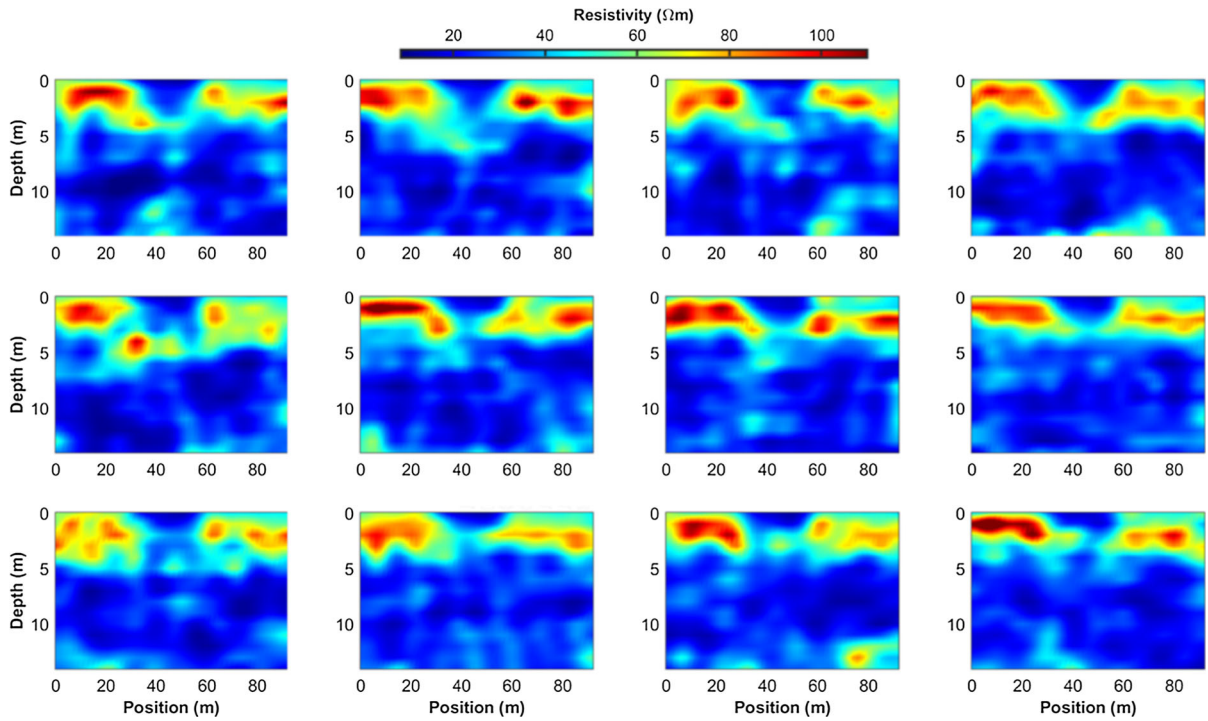


Figure 17
Some examples of models sampled at the third iteration for the test with 1000 models

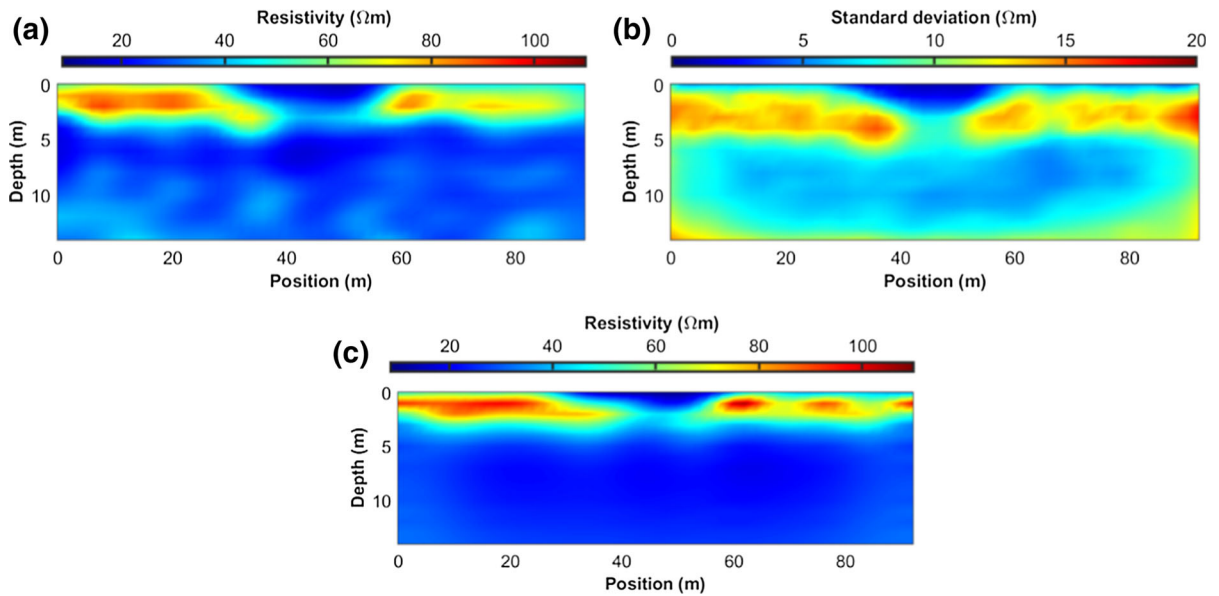


Figure 18

a Posterior mean estimated by the DEMC algorithm. **b** Posterior standard deviation estimated by the DEMC method. **c** Final model provided by the deterministic inversion

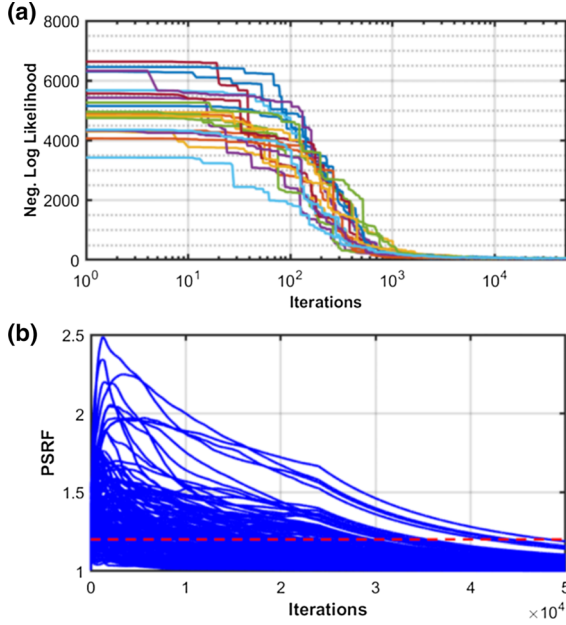


Figure 19

a Evolution of the negative log-likelihood for the 20 chains during the DEMC inversion. Each color represents a chain. **b** Evolution of the PSRF for the model parameters during the DEMC inversion. The dotted red line at 1.2 represents the threshold of convergence

part of the subsurface is characterized by the lowest ambiguity while the uncertainties increase within the high resistivity formation, and at the lateral edges and the deepest part of the model due to the lower illumination. The computing times listed in Table 3 still refer to parallel Matlab codes running on the hardware configuration previously described. All tests yielded final mean models producing similar good matches with the observed data.

Figure 15 compares the evolution of the L2 norm data misfit for the tests employing ensembles of 200, 500, and 1000 models. Similar to the synthetic experiments we observe that just three iterations are needed to attain a stable misfit value and that a larger ensemble guarantees a faster decrease of the error in the very first iterations. From the previous considerations it emerges that in this case, the ensemble with 1000 models is the one that guarantees the best compromise between the reliability of the results and the computational effort. This test attains convergence in 16 min.

Figure 16 shows some examples of resistivity sections sampled from the prior and pertaining to the

test with 1000 models. These resistivity sections evolve into those shown in Fig. 17, where we observe that all the considered models univocally predict lateral high resistivity bodies located around 2 m depth.

To assess the quality of the results we compare the outcomes provided by the EB-ERT inversion with those achieved by a DEMC algorithm sampling the same reduced space, and with a standard deterministic least-squares inversion running in the full data and model spaces. The DEMC employed 20 chains evolving for 50,000 iterations and takes 7 days, approximately, of computing time to converge. Figure 18a, b represent the posterior mean models and standard deviation provided by the DEMC algorithm, whereas Fig. 18c shows the predictions of the deterministic inversion. The close agreement between these results and those achieved by the EB-ERT approach (Fig. 14d) ensures us about the applicability and the reliability of the proposed inversion approach. Similar to the synthetic example, we observe that the EB approach slightly underestimates the uncertainties for the model parameters poorly informed by the data. The evolution of the negative log-likelihood value and of the potential scale reduction factor for the DEMC inversion (Fig. 19), highlights that also in the reduced domain the stationary regime is reached in more than 5000 iterations, while more than 45,000 iterations are needed to achieve stable posterior assessments. Note that a much longer sampling stage and longer burn-in period would have been needed to sample the PPD in the full model space.

These results demonstrate that the EB-ERT provides estimations comparable with those provided by the DEMC with a reduction of the total computing time of more than 3 orders of magnitudes. The deterministic solution is obtained in less than a minute, but this method does not provide neither an estimate of the uncertainty affecting the recovered model nor an ensemble of models that reproduces the acquired data.

In this field experiment, all the methods provide final model estimations that well reproduce the observed pseudosection (Fig. 20). In Table 4 we observe that the RMSE for the data is similar for EB and MCMC and slightly lower for the deterministic

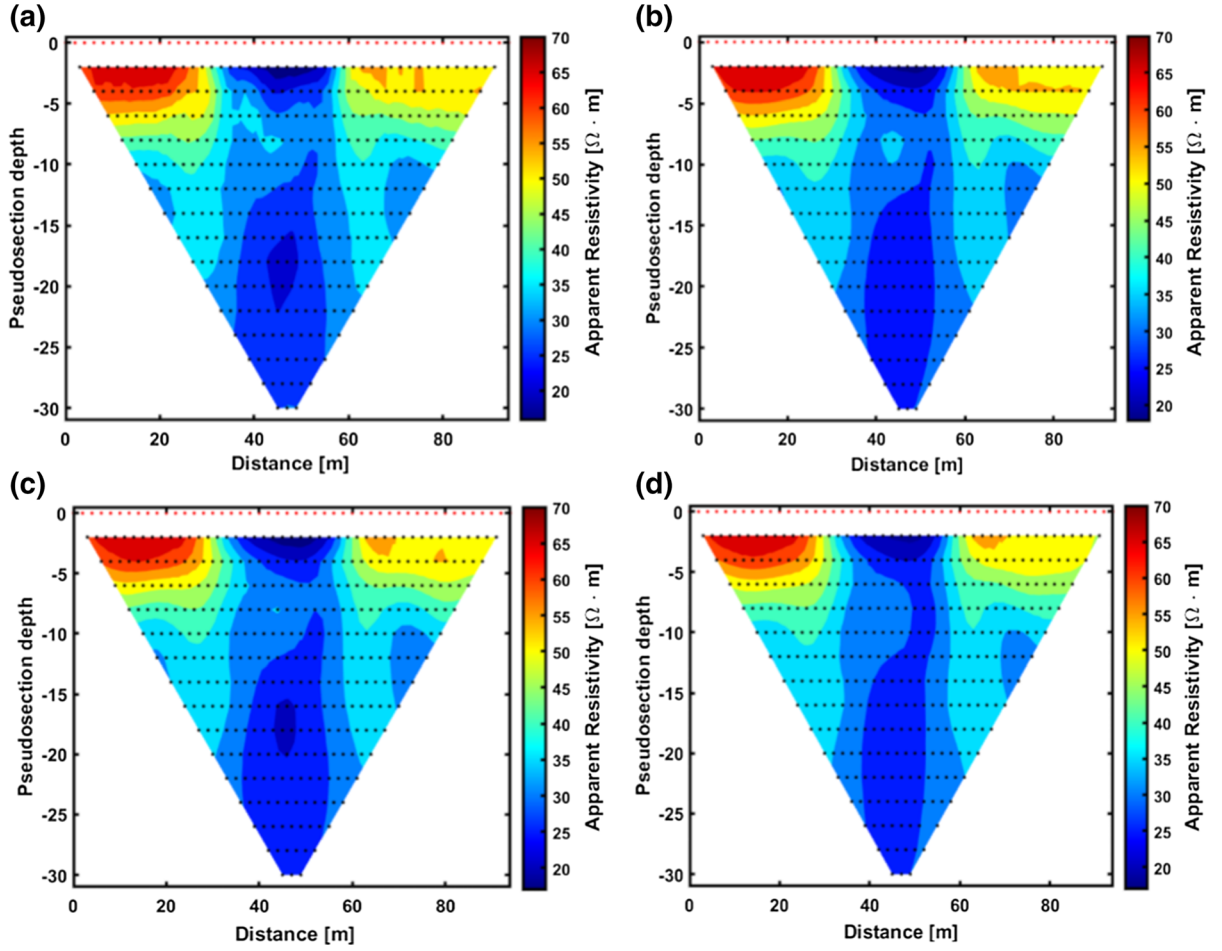


Figure 20

a Observed apparent resistivity pseudosection. Calculated apparent resistivity pseudosections from, **b** the mean model estimated by the EB-ERT inversion with 1000 ensemble members (see Fig. 14d), **c** the deterministic solution (see Fig. 18c), **d** the mean model estimated by the MCMC inversion (see Fig. 18a)

Table 4

Root-mean-square data errors for the EB, MCMC, and deterministic inversion

	EB	MCMC	Deterministic
RMSE data	2.32	2.44	2.11

inversion. However, note that the predicted data associated with the deterministic approach is computed on the model that minimizes the error function, while for the EB and MCMC inversion we compute the predicted data on posterior mean models

computed taking into account the entire ensemble (for EB) and all the samples collected after the burn-in period (for MCMC).

5. Conclusions

In this paper, we applied an ensemble-based approach to cast the ERT inversion (EB-ERT) into a Bayesian setting, where prior assumptions and reduced model, and data reparametrizations are incorporated to regularize the problem. The approach uses the ES-MDA stochastic algorithm that provides multiple realizations of the subsurface resistivity

values that honor the observed data and that approximates the target posterior probability distribution. The computational cost of the implemented approach is mostly dependent on the number of ensemble members and the cost of running the forward model. The DCT compression of the data and model space mitigated the ill-conditioning of the ERT problem and also allowed running the inversion in lower-dimensional data and model spaces. This strategy consequently reduced the computational cost of the EB-ERT inversion because reliable model estimations and uncertainty assessments can be obtained with smaller ensembles.

Inversions of both synthetic and field data illustrated the applicability and the reliability of the proposed inversion approach. We demonstrated that in the context of ERT application, the implemented inversion provides final mean models and uncertainty quantifications comparable to those yielded by a much more computationally expensive MCMC sampling, although the EB tends to slightly underpredict the uncertainties associated with the parameters poorly informed by the data. In both the synthetic and field data application, the EB-ERT approach reduced the computational effort of the Bayesian ERT inversion by more than 3 orders of magnitude with respect to the MCMC algorithm. The main advantage of probabilistic methods over standard deterministic inversions is that they provide multiple stochastic realizations that can be visualized and queried to provide meaningful information about the accuracy and the precision of the recovered solution.

The implemented approach can be applied to other nonlinear inverse problems, and it provides uncertainty estimations without requiring the linearization of the forward operator, thus avoiding the modeling errors introduced by the linear approximation. In this work, we generated the initial ensemble using the simple kriging approach, but other geostatistical methods, such as variogram-based and multiple-point geostatistics algorithms, can be included in the inversion framework for generating the initial ensemble.

Acknowledgements

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. The monitoring system (G.Re.T.A.) was developed and installed in Colorno pilot site with technical support of LSI-Lastem S.r.l. and scientific support of Politecnico di Milano.

Funding

Open access funding provided by Università di Pisa within the CRUI-CARE Agreement. 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.

Data Availability

Data available on request from the authors.

Code Availability

Code available on request from the authors.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- Aleardi, M. (2019). Using orthogonal Legendre polynomials to parameterize global geophysical optimizations: Applications to seismic-petrophysical inversion and 1D elastic full-waveform inversion. *Geophysical Prospecting*, 67(2), 331–348.
- Aleardi, M. (2020). Discrete cosine transform for parameter space reduction in linear and non-linear AVA inversions. *Journal of Applied Geophysics*, 179, 104106.
- Aleardi, M., Ciabbari, F., & 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.
- Aleardi, M., & Salusti, A. (2020). 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., & Salusti, A. (2021). Elastic pre-stack seismic inversion through discrete cosine transform reparameterization and convolutional neural networks. *Geophysics*, 86(1), 1JF-V89.
- Aleardi, M., Vinciguerra, A., & Hojat, A. (2020). A geostatistical markov chain monte carlo inversion algorithm for electrical resistivity tomography. *Near Surface Geophysics*. <https://doi.org/10.1002/nsg.12133>
- Arosio, D., Munda, S., Tresoldi, G., Papini, M., Longoni, L., & Zanzi, L. (2017). A customized resistivity system for monitoring saturation and seepage in earthen levees: Installation and validation. *Open Geosciences*, 9(1), 457–467. <https://doi.org/10.1515/geo-2017-0035>
- Aster, R. C., Borchers, B., & Thurber, C. H. (2018). *Parameter estimation and inverse problems*. Elsevier.
- Bièvre, G., Oxarango, L., Günther, T., Goutaland, D., & Massardi, M. (2018). Improvement of 2D ERT measurements conducted along a small earth-filled dyke using 3D topographic data and 3D computation of geometric factors. *Journal of Applied Geophysics*, 153, 100–112.
- Britanak, V., Yip, P. C., & Rao, K. R. (2010). *Discrete cosine and sine transforms: General properties, fast algorithms and integer approximations*. Elsevier.
- Brooks, S. P., & Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical STATISTICS*, 7(4), 434–455.
- Chambers, J. E., Kuras, O., Meldrum, P. I., Ogilvy, R. D., & Hollands, J. (2006). Electrical resistivity tomography applied to geologic, hydrogeologic, and engineering investigations at a former waste-disposal site. *Geophysics*, 71(6), B231–B239. <https://doi.org/10.1190/1.2360184>
- Dahlin, T. (2020). Geoelectrical monitoring of embankment dams for detection of anomalous seepage and internal erosion—Experiences and work in progress in Sweden. In *Fifth International Conference on Engineering Geophysics (ICEG)*, Al Ain, UAE. <https://doi.org/10.1190/iceg2019-053.1>.
- Dejtrakulwong, P., Mukerji, T., & Mavko, G. (2012). Using kernel principal component analysis to interpret seismic signatures of thin shaly-sand reservoirs. In *SEG Technical Program Expanded Abstracts*. Society of Exploration Geophysicists, pp. 1–5.
- Emerick, A. A., & Reynolds, A. C. (2013). Ensemble smoother with multiple data assimilation. *Computers and Geosciences*, 55, 3–15.
- Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5), 10143–10162.
- Evensen, G. (2009). *Data assimilation: The ensemble Kalman filter*. Springer.
- Fernández-Martínez, J. L., Fernández-Muñiz, Z., Pallero, J. L. G., & Bonvalot, S. (2017). Linear geophysical inversion via the discrete cosine pseudo-inverse: Application to potential fields. *Geophysical Prospecting*, 65, 94–111.
- Grana, D., Passos de Figueiredo, L., & Azevedo, L. (2019). Uncertainty quantification in Bayesian inverse problems with model and data dimension reduction. *Geophysics*, 84(6), M15–M24.
- Hermans, T., & 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), e2019GL85877.
- Hojat, A., Arosio, D., Ivanov, V. I., Loke, M. H., Longoni, L., Papini, M., Tresoldi, G., & 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 Geophysics*, 18, 427–443. <https://doi.org/10.1002/nsg.12110>
- Hojat, A., Arosio, D., Ivanov, V. I., Longoni, L., Papini, M., Scaioni, M., Tresoldi, G., & Zanzi, L. (2019a). Geoelectrical characterization and monitoring of slopes on a rainfall-triggered landslide simulator. *Journal of Applied Geophysics*, 170, 103844. <https://doi.org/10.1016/j.jappgeo.2019.103844>
- Hojat, A., Arosio, D., Longoni, L., Papini, M., Tresoldi, G., & Zanzi, L. (2019b). Installation and validation of a customized resistivity system for permanent monitoring of a river embankment. In *EAGE-GSM 2nd Asia Pacific Meeting on Near Surface Geoscience and Engineering*, Kuala Lumpur, Malaysia. <https://doi.org/10.3997/2214-4609.201900421>.
- Jin, L., Sen, M. K., & Stoffa, P. L. (2008). One-dimensional pre-stack seismic waveform inversion Using Ensemble Kalman Filter. In *SEG Technical Program Expanded Abstracts*. Society of Exploration Geophysicists, pp. 1920–1924.
- Karaoulis, M., Revil, A., Tsourlos, P., Werkema, D. D., & Minsley, B. J. (2013). IP4DI: A software for time-lapse 2D/3D DC-resistivity and induced polarization tomography. *Computers and Geosciences*, 54, 164–170.
- Karaoulis, M., Tsourlos, P., Kim, J. H., & Revil, A. (2014). 4D time-lapse ERT inversion: Introducing combined time and space constraints. *Near Surface Geophysics*, 12(1), 25–34.
- Karimi-Nasab, S., Hojat, A., Kamkar-Rouhani, A., Akbari-Javar, H., & Maknooni, S. (2011). Successful use of geoelectrical surveys in area no. 3 of the Gol-e-Gohar iron ore mine Iran. *Journal of Mine Water and the Environment*, 30(3), 208–215. <https://doi.org/10.1007/s10230-011-0135-7>
- Li, Y., & Oldenburg, D. W. (2003). Fast inversion of large-scale magnetic data using wavelet transforms and a logarithmic barrier method. *Geophysical Journal International*, 152(2), 251–265.

- Liu, M., & Grana, D. (2018a). Ensemble-based seismic history matching with data reparameterization using convolutional autoencoder. In *SEG Technical Program Expanded Abstracts*. Society of Exploration Geophysicists, pp. 3156–3160.
- Liu, M., & Grana, D. (2018b). Stochastic nonlinear inversion of seismic data for the estimation of petroelastic properties using the ensemble smoother and data reparameterization. *Geophysics*, 83(3), M25–M39.
- Liu, M., & Grana, D. (2018c). Ensemble-based joint inversion of PP and PS seismic data using full Zoeppritz equations. In *SEG Technical Program Expanded Abstracts*. Society of Exploration Geophysicists, pp. 511–515.
- Lochbühler, T., Breen, S. J., Detwiler, R. L., Vrugt, J. A., & Linde, N. (2014). Probabilistic electrical resistivity tomography of a CO₂ sequestration analog. *Journal of Applied Geophysics*, 107, 80–92.
- Loke, M. H., Papadopoulos, N., Wilkinson, P. B., Oikonomou, D., Simyrdanis, K., & Rucker, D. F. (2020). The inversion of data from very large three-dimensional electrical resistivity tomography mobile surveys. *Geophysical Prospecting*, 68(8), 2579–2597.
- Menke, W. (2018). *Geophysical data analysis: Discrete inverse theory*. Academic Press.
- Moradipour, M., Ranjbar, H., Hojat, A., Karimi-Nasab, S., & Daneshpajouh, S. (2016). Laboratory and field measurements of electrical resistivity to study heap leaching pad no. 3 at Sarcheshmeh copper mine. In *22nd European Meeting of Environmental and Engineering Geophysics*. <https://doi.org/10.3997/2214-4609.201602140>.
- Müller, K., Vanderborght, J., Englert, A., Kemna, A., Huisman, J. A., Rings, J., & 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*. <https://doi.org/10.1029/2008WR007595>
- Pidlisecky, A., & Knight, R. (2008). FW2_5D: A MATLAB 2.5-D electrical resistivity modeling code. *Computers and Geosciences*, 34(12), 1645–1654.
- Pollock, D., & Cirpka, O. A. (2012). Fully coupled hydrogeophysical inversion of a laboratory salt tracer experiment monitored by Electrical resistivity tomography. *Water Resources Research*. <https://doi.org/10.1029/2011WR010779>
- Pradhan, A., & Mukerji, T. (2020). Seismic Bayesian evidential learning: Estimation and uncertainty quantification of sub-resolution reservoir properties. *Computational Geosciences*, 24, 1121–1140.
- Sajeva, A., Aleardi, M., Mazzotti, A., Bienati, N., & Stucchi, E. (2014). Estimation of velocity macro-models using stochastic full-waveform inversion. In *SEG Technical Program Expanded Abstracts 2014*. Society of Exploration Geophysicists, p.1227–1231.
- Sambridge, M., & Mosegaard, K. (2002). Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40(3), 3–1
- Sen, M. K., & Stoffa, P. L. (2013). *Global optimization methods in geophysical inversion*. Cambridge University Press.
- Supper, R., Ottowitz, D., Jochum, B., Kim, J. H., Romer, A., Baron, I., Pfeiler, S., Lovisolo, M., Gruber, S., & Vecchiotti, F. (2014). Geoelectrical monitoring: An innovative method to supplement landslide surveillance and early warning. *Near Surface Geophysics*, 12(1), 133–150. <https://doi.org/10.3997/1873-0604.2013060>
- Tarantola, A. (2005). *Inverse problem theory*. SIAM.
- Thurin, J., Brossier, R., & Métivier, L. (2019). Ensemble-based uncertainty estimation in full waveform inversion. *Geophysical Journal International*, 219(3), 1613–1635
- Tresoldi, G., Arosio, D., Hojat, A., Longoni, L., Papini, M., & Zanzi, L. (2019). Long-term hydrogeophysical monitoring of the internal conditions of river levees. *Engineering Geology*, 259, 105139. <https://doi.org/10.1016/j.enggeo.2019.05.016>
- Tveit, S., Bakr, S. A., Lien, M., & Mannseth, T. (2015). Ensemble-based Bayesian inversion of CSEM data for subsurface structure identification. *Geophysical Journal International*, 201(3), 1849–1867
- Tveit, S., Mannseth, T., Park, J., Sauvin, G., & Agersborg, R. (2020). Combining CSEM or gravity inversion with seismic AVO inversion, with application to monitoring of large-scale CO₂ injection. *Computational Geosciences*, 24, 1–20
- Vrugt, J. A. (2016). Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling and Software*, 75, 273–316
- Whiteley, J., Chambers, J.E., & 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*, pp. 27.
- Zhang, Y., Ghodrati, A., & 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