Geophysical Prospecting
EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS
Geophysical Prospecting, 2021 doi: 10.1111/1365-2478.13148

Discrete cosine transform for parameter space reduction in Bayesian electrical resistivity tomography

A. Vinciguerra[1,2]*, M. Aleardi[2], A. Hojat[3,4], M. H. Loke[5] and E. Stucchi[2]

[1]*Earth Sciences Department, University of Florence, via G. La Pira 4, Florence, 50121, Italy,* [2]*Earth Sciences Department, University of Pisa, via S. Maria 53, Pisa, 56126, Italy,* [3]*Department of Mining Engineering, Shahid Bahonar University of Kerman, Kerman, 76188, Iran,* [4]*Department of Civil and Environmental Engineering, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano, 20133, Italy, and* [5]*Geotomo Software Sdn Bhd, Gelugor, Penang 11700, Malaysia*

Received March 2021, revision accepted August 2021

ABSTRACT

Electrical resistivity tomography is a non-linear and ill-posed geophysical inverse problem that is usually solved through gradient-descent methods. This strategy is computationally fast and easy to implement but impedes accurate uncertainty appraisals. We present a probabilistic approach to two-dimensional electrical resistivity tomography in which a Markov chain Monte Carlo algorithm is used to numerically evaluate the posterior probability density function that fully quantifies the uncertainty affecting the recovered solution. The main drawback of Markov chain Monte Carlo approaches is related to the considerable number of sampled models needed to achieve accurate posterior assessments in high-dimensional parameter spaces. Therefore, to reduce the computational burden of the inversion process, we employ the differential evolution Markov chain, a hybrid method between non-linear optimization and Markov chain Monte Carlo sampling, which exploits multiple and interactive chains to speed up the probabilistic sampling. Moreover, the discrete cosine transform reparameterization is employed to reduce the dimensionality of the parameter space removing the high-frequency components of the resistivity model which are not sensitive to data. In this framework, the unknown parameters become the series of coefficients associated with the retained discrete cosine transform basis functions. First, synthetic data inversions are used to validate the proposed method and to demonstrate the benefits provided by the discrete cosine transform compression. To this end, we compare the outcomes of the implemented approach with those provided by a differential evolution Markov chain algorithm running in the full, un-reduced model space. Then, we apply the method to invert field data acquired along a river embankment. The results yielded by the implemented approach are also benchmarked against a standard local inversion algorithm. The proposed Bayesian inversion provides posterior mean models in agreement with the predictions achieved by the gradient-based inversion, but it also provides model uncertainties, which can be used for penetration depth and resolution limit identification.

Key words: 2D resistivity, Bayesian, Inversion.

INTRODUCTION

Geoelectrical methods are commonly applied for mining exploration, geotechnical characterization and hydrological studies (Pelton *et al.*, 1976; Sumner, 1976; Goldie, 2002;

*Email: alessandro.vinciguerra@unifi.it

Loke *et al.*, 2013). Electrical resistivity tomography (ERT) is usually applied for groundwater exploration, geotechnical characterization, mapping of contaminant plumes, landfill and levees monitoring (Uhlemann *et al.*, 2017; Tresoldi *et al.*, 2018, 2019; Hojat *et al.*, 2019; Helene *et al.*, 2020). Typically, deterministic gradient-based algorithms (Zhang *et al.*, 2005; Pidlisecky and Knight, 2008) are employed to tackle the ERT problem. These methods linearize the problem around an initial solution, thereby losing the information for accurate uncertainty appraisals. Moreover, these algorithms are prone to get trapped in local minima of the misfit function (e.g. the L2 norm difference between observed and predicted data) and thus the starting model must lie within the valley of attraction of the global minimum. In addition, the ERT inversion is an ill-conditioned inverse problem in which many subsurface models equally fit the measured apparent resistivity pseudo-section. Therefore, an accurate estimation of the uncertainties affecting the retrieved solution would be of great help in the interpretation phase because it could be used to generate multiple subsurface scenarios all in accordance with the observed data. The Bayesian approach is commonly employed to cast an inverse problem into a solid probabilistic framework. In this context, the final solution of the inversion is the so-called posterior probability density (PPD) function in model space (Ramirez *et al.*, 2005; Tarantola, 2005; Sen and Stoffa, 2013; Aleardi *et al.*, 2018; Galetti and Curtis, 2018; Aleardi and Salusti, 2020) that fully quantifies the ambiguities in the recovered model. However, the ERT is a non-linear problem and for this reason the PPD cannot be expressed in a closed form, but it must be numerically evaluated, for example by employing Markov Chain Monte Carlo algorithms (MCMC; Sambridge and Mosegaard, 2000). These algorithms transform the Bayesian inversion process into a sampling problem in which the sampling density is proportional to the PPD. Although the increasing computational power provided by modern parallel architectures has considerably encouraged the applications of MCMC methods to solve geophysical problems (Fichtner and Zunino *et al.*, 2019; Stuart *et al.*, 2019; Aleardi *et al.*, 2020; Aleardi, 2020a), 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 exchange models slowly when several areas of the model space are equally probable, that is when the target distribution is multimodal (Holmes *et al.*, 2017; Scalzo *et al.*, 2019). Besides, the sampling ability of MCMC algorithms severely decreases in highly dimensional model spaces due to the so-called curse of the dimensionality problem (Curtis *et al.*, 2001). Over the past decades, many MCMC approaches have been proposed to mitigate this issue, for example, a simple strategy is to use 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, it turns out to be inefficient in high-dimensional problems where the curse of dimensionality makes the target distribution highly localized within each model space dimension, and, consequently, the MCMC chains are likely to get trapped in the local maxima of the PPD. There have been many attempts to improve the sampling ability of MCMC algorithms in high-dimensional spaces, for example, by hybridizing standard MCMC algorithms with global search methods (e.g. differential evolution Markov chain "DEMC" or differential evolution adaptive Metropolis; Turner *et al.*, 2013; Vrugt, 2016).

Another viable strategy to mitigate the curse of dimensionality and to reduce the computational complexity of high-dimensional inverse problems is to compress the model space through appropriate reparameterization techniques (Fernández-Martínez *et al.*, 2011; Azevedo *et al.*, 2016; Aleardi, 2019; Szabó and Dobróka, 2019; Numes *et al.*, 2019; Aleardi 2020b). However, it should be noted that the parameterization of an inverse problem must always constitute a compromise between model resolution and model uncertainty (Grana *et al.*, 2019).

In this work, we use the discrete cosine transform (DCT; Ahmed *et al.*, 1974) to reparameterize the Bayesian ERT inversion solved through a DEMC sampling of the parameter space. This transform, originally developed for signal processing and imaging compression, is comparable to the Fourier transform, but it uses only cosines as bases functions to signal reconstruction so that the computed coefficients are real numbers. Therefore, the DCT of a signal (i.e. expressing the subsurface resistivity model) indicates the energy distribution of the signal in the frequency domain spectrum. Usually, most of the energy of the signal is expressed by low-order DCT coefficients. Consequently, this mathematical transformation can be used for model compression, which is accomplished by setting the coefficients of the base function terms beyond a certain threshold equal to zero (such as a low pass filtering). In this context, the coefficients associated with the retained bases become the parameters that express the signal, and in the inverse problem, they are the unknown parameters to retrieve.

We first discuss some synthetic examples in which the observed data have been derived from a schematic subsurface model through a 2.5D Finite Elements forward modelling code taken from the boundless electrical resistivity

tomography open-source software (Günther *et al.*, 2006). In this section, to demonstrate the benefits provided by the DCT reparameterization of DEMC-ERT inversion, we compare the outcomes provided by the proposed approach with those yielded by a DEMC inversion running in the unreduced model space. Then, we discuss the application to field data acquired along a river embankment. In both the synthetic and field examples, the results of the implemented algorithm are benchmarked with the predictions of a gradient-based inversion approach (Loke, 2013).

METHODS

In this section, we first describe the discrete cosine transform (DCT) before discussing the implemented differential evolution Markov chain and electrical resistivity tomography inversion.

The discrete cosine transform

The DCT is a linear orthogonal transformation that decomposes a signal into a sum of cosine functions oscillating at different frequencies. This transformation has successfully been used in imaging compression standards due to its energy compaction properties (Jain, 1989). Considering that signals are real in most applications, the DCT can be used instead of the discrete Fourier transform to reduce the calculation redundancy. There are several variants of DCT formulation, but the most employed is the DCT-II (Ahmed *et al.*, 1974). For simplicity from here on we refer to DCT-II as the DCT. Equation (1) represents one-dimensional (1D) DCT equation, and in Figure 1 we illustrate an example of five cosinuisodal 1D bases:

$$\begin{cases} R\ (k_x) = \dfrac{1}{\sqrt{M_x}} \displaystyle\sum_{x\,=\,0}^{M_x-1} \boldsymbol{\rho}\,(x)\,, \quad \text{if } k_x = 0 \\ R\ (k_x) = \sqrt{\dfrac{2}{M_x}} \displaystyle\sum_{x\,=\,0}^{M_x-1} \boldsymbol{\rho}\,(x) \cos\left(\dfrac{(2x+1)\,\pi k_x}{2M_x}\right), \quad \text{if } k_x \neq 0 \end{cases}, \tag{1}$$

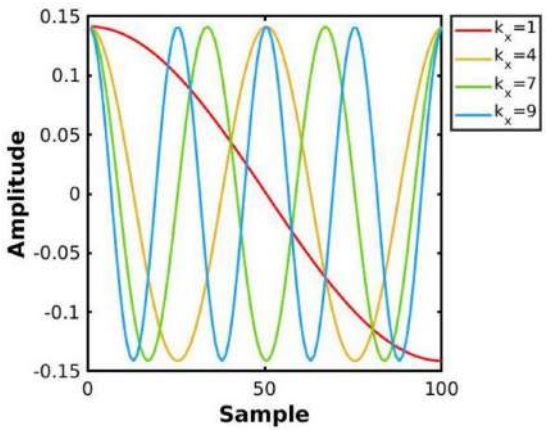

Figure 1 Example of five DCT bases associated with DCT coefficients with different orders.

where $\mathbf{R}$ corresponds to the M_x coefficients that describe the original 1D signal $\boldsymbol{\rho}(x)$ of length M_x and k_x represents the order of each DCT coefficient.

The DCT can also be extended to multi-dimensional signals (i.e. two-dimensional [2D] matrices), and such multi-dimensional transformation follows straightforwardly from the 1D definition because it is simply a separable product (equivalently, a composition) of DCTs along each dimension. For example, a 2D DCT transformation of a resistivity model $\rho(x, y)$ of $[M_y, M_x]$ can be computed as follows:

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

where $\mathbf{R}(k_x, k_y)$ represents the transformed resistivity model, whereas x and y indicate the horizontal and vertical coordinate axes, respectively; M_x and M_y denote the number of parameters along the x and y directions, respectively. Equation (2) can be compactly rearranged in the matrix form:

$$\mathbf{R} = \mathbf{B}_y\, \boldsymbol{\rho} \mathbf{B}_x^T, \tag{3}$$

where $\mathbf{B}_x$ and $\mathbf{B}_y$ are matrices with dimensions $M_x \times M_x$ and $M_y \times M_y$, respectively, and contain the orthogonal DCT basis functions, whereas the $M_y \times M_x$ matrix $\mathbf{R}$ contains the DCT coefficients. Most of the spatial variability of the resistivity model is explained by low-order DCT coefficients, and for this

reason an approximation of the subsurface resistivity model can be obtained as follows:

$$\bar{\rho} = \left(\mathbf{B}_y^q\right)^T \mathbf{R}_{qp}\mathbf{B}_x^p, \tag{4}$$

where $\bar{\rho}$ 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 scalars q and p represent the retained number of base functions along the y and x directions used to derive the approximated resistivity model. Therefore, the DCT transformation allows for a reduction of the $(M_y \times M_x)-$2D full resistivity model space to a $(q \times p)$ 2D DCT compressed parameter space with $p < M_x$ and $q < M_y$. In all the following inversion tests, we assume log-Gaussian prior distributions for the resistivity, and given the linearity of the DCT transformation, the assumed a priori mean vectors and a priori covariance matrices can be analytically projected onto the DCT space (see Aleardi, 2020b). For example, let $\mathbf{C}_m$ be the prior model covariance in the resistivity space. Then, the prior covariance in the DCT space ($\mathbf{C}_r$) can be obtained as follows:

$$\mathbf{C}_r = \left(\mathbf{B}_y^q\right)^T \mathbf{C}_m\mathbf{B}_x^p, \tag{5}$$

In our application, the $\mathbf{C}_m$ matrix also codes a 2D stationary and Gaussian variogram model that expresses the assumed spatial variability of the resistivity values. The correlation function of the resistivity model is expressed, considering, for example the y-direction, by the following function:

$$\boldsymbol{\tau}_y = \exp\left(-\frac{h_y^2}{a_y^2}\right), \tag{6}$$

where h_y is the spatial distance of the autocorrelation function along the y-direction and a_y is the effective range of the variogram along the y-direction. The model covariance matrix $\mathbf{C}_m$ is computed as the double Kronecker product between the prior variance, the spatial correlation function τ_y and the spatial correlation τ_x (along the x-direction):

$$\mathbf{C}_m = ((\mathrm{var}(\boldsymbol{\rho}) \otimes \mathbf{T}(\boldsymbol{\tau}_x)) \otimes \mathbf{T}(\boldsymbol{\tau}_y)), \tag{7}$$

where $\mathrm{var}(\boldsymbol{\rho})$ is the variance of the resistivity model, $\boldsymbol{\tau}_y$ the correlation function along the y-direction (Equation 6), $\mathbf{T}$ indicates the Toeplitz matrix and $\otimes$ stands for the Kronecker product. As previously mentioned, there exists a trade-off between model resolution and the number of coefficients employed (i.e. the spatial resolution of the recovered model increases as the number of retained DCT coefficients increases). For this reason, after selecting the prior model for the Bayesian inversion, the next step involves the estimation of the optimal number of DCT coefficients needed to approximate the subsurface model. To this end, we quantify how the variability of the resistivity model drawn from the prior changes as the number of DCT basis functions varies. We compute the variability as the ratio between the variance of the approximated model and the variance of the uncompressed model. In our work, the retained number of coefficients preserves 98% of the variability. However, there exists a different strategy to select the DCT coefficients to retain without fixing the value of q and p (Mogadhas *et al.*, 2019), but we apply the simplest approach.

Markov chain Monte Carlo inversion and the differential evolution Markov chain

The final solution of a geophysical Bayesian inversion is the posterior probability density (PPD) function that expresses the probability of the model $\mathbf{m}$ conditioned upon the observed data $\mathbf{d}$:

$$p(\mathbf{m}|\mathbf{d}) = \frac{p(\mathbf{d}|\mathbf{m})\, p(\mathbf{m})}{p(\mathbf{d})}, \tag{8}$$

where $p(\mathbf{m}|\mathbf{d})$ is the target density, $p(\mathbf{m})$ and $p(\mathbf{d})$ are the prior distributions of model parameters and data, respectively; $p(\mathbf{d}|\mathbf{m})$ is the so-called likelihood function, which under the assumption of uncorrelated and normally distributed data, takes the following form:

$$p(\mathbf{d}|\mathbf{m}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-0.5\frac{\left(\mathbf{d}_i^{\mathrm{pre}}(\mathbf{m}) - \mathbf{d}_i\right)^2}{\sigma_i^2}\right], \tag{9}$$

where $\mathbf{d}_i^{\mathrm{pre}}$ denotes the ith predicted data, $\mathbf{d}_i$ is the ith observed data and σ_i^2 represents the variance of the noise associated with the ith data point. The target distribution can be numerically sampled by adopting the Metropolis–Hasting rule (Metropolis *et al.*, 1953; Hastings, 1970), which defines the probability to move from the current state of the Markov chain $\mathbf{m}$ to the proposed (perturbed) state $\mathbf{m}'$ as follows:

$$\alpha = p(\mathbf{m}'|\mathbf{m}) = \min\left[1, \frac{p(\mathbf{m}')}{p(\mathbf{m})} \times \frac{p(\mathbf{d}|\mathbf{m}')}{p(\mathbf{d}|\mathbf{m})} \times \frac{q(\mathbf{m}|\mathbf{m}')}{q(\mathbf{m}'|\mathbf{m})}\right], \tag{10}$$

where q is the proposal distribution that defines the new state $\mathbf{m}'$ as a random deviate from a probability distribution $q(\mathbf{m}'|\mathbf{m})$ conditioned only on the current state $\mathbf{m}$. The proposal ratio term becomes unity if symmetric proposals (for instance, Gaussian proposal) are used. If $\mathbf{m}'$ is accepted $\mathbf{m} = \mathbf{m}'$. Otherwise, $\mathbf{m}$ is repeated in the chain, and another model is generated as a random deviate from $\mathbf{m}$. The

ensemble of sampled models after the burn-in period is used to numerically compute the statistical properties (e.g. mean, mode, standard deviations, marginal densities) of the PPD.

Theoretically, for an infinite number of sampled models, the PPD estimated by an MCMC algorithm does not depend on the choice of the proposal. However, from a more practical perspective, the efficiency of MCMC methods increases if the proposal distribution is a good approximation of the target distribution. For this reason, the definition of an appropriate proposal distribution is crucial for an efficient probabilistic sampling: suboptimal choices of the proposal often result in a persistent rejection of models or entrapments in local optima of the PPD (Vrugt, 2016). To partially attenuate this issue, the PPD can be numerically evaluated from the models collected by multiple and independent MCMC chains. It has been demonstrated that a mixing of the information (i.e. sampled models) brought by the different chains, considerably increases the efficiency and the rate of convergence of MCMC algorithms (Craiu *et al.*, 2005; Vrugt, 2016). One of the most popular MCMC algorithms that takes advantage of multiple and interactive chains is the differential evolution Markov chain (DEMC). It exploits some principles coming from the genetic algorithms for population evolution, whereas it uses a Metropolis selection rule (Equation 10) to decide whether candidate states should replace their parents or not (Ter Braak, 2006). In DEMC, multiple Markov chains and multivariate proposals are generated on the fly from the collection of chains using differential evolution principles. If the state of a single chain is given by the p-vector $\mathbf{m}$ (proposal model), then at each iteration $t-1$ the Q chains in DEMC define a population $\mathrm{M} = \mathbf{m}_{t-1}^1, \mathbf{m}_{t-1}^2, \ldots, \mathbf{m}_{t-1}^Q$, which corresponds to a $Q \times p$ matrix, with each chain as a row. Then multivariate proposal $\mathbf{m}_p$ is defined for each chain:

$$\mathbf{m}_p^i = \mathbf{m}_{t-1}^i + \gamma\left(\mathbf{m}_{t-1}^a - \mathbf{m}_{t-1}^b\right) + \epsilon, a \neq b \neq i, \tag{11}$$

where i is the index of the current chain, γ denotes the jump rate, a and b are integer values drawn from {1,..., $i-1$, $i+1$,..., Q}, and ϵ is 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 according to equation (10). If the proposal $\mathbf{m}_p^i$ 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 is usually set to $\gamma = 2.38/2p$ (Vrugt, 2016). Besides, to promote mode-jumping there exists a 10% probability for $\gamma = 1$, which is a significant strength of DEMC compared with more standard MCMC methods (i.e. random walk Metropolis). Also, note that the DEMC avoid for the user the task of the selection of the proposal distribution because the random deviate from the current model is automatically determined as $\gamma(\mathbf{m}_{t-1}^a - \mathbf{m}_{t-1}^b)$ (Equation 11).

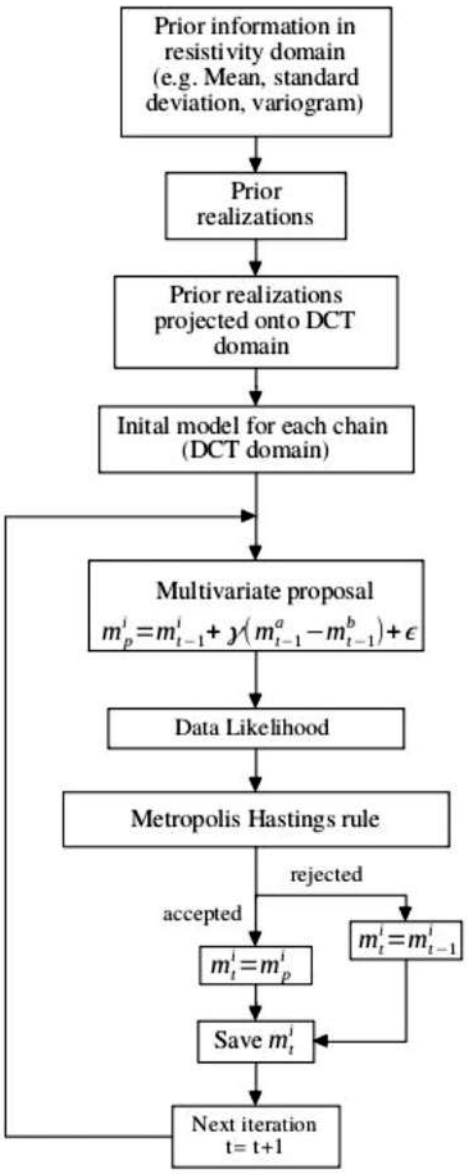

Figure 2 Flow chart of the procedures for DCT-DEMC inversion.

If the MCMC inversions run in the reduced DCT space, the sampled models must be projected back onto the resistivity space (see equation (3)) just before the forward modelling phase that gives the predicted data needed to compute the likelihood value. The posterior model can be numerically derived from the ensemble of DCT models collected during the sampling stage, after projection onto the resistivity space. The flow chart of the procedure is shown in Figure 2.

Any MCMC algorithm must be run until a stable estimation of the PPD has been attained. The convergence of the MCMC sampling can be monitored, for example, by comparing the PPDs estimated by each chain in the first and second halves of the sampling stage. If these PPDs are in good agreement, no further sampling is needed. A more reliable measure of convergence is given by the potential-scale-reduction-factor (PSRF) that compares the difference between the 'within-chain' and 'between-chain' estimated variances for each model parameter (for details, see Gelman *et al*., 1995; Brooks *et al*., 1998). The PSRF decreases to 1 as the number of drawn samples tends to be infinite. A high PSRF value indicates that the variance within the walks is small compared to that between the walks and that a long walk is needed to attain convergence. Usually, a PSRF lower than 1.2 for a given unknown indicates that convergence has been achieved for that model parameter.

RESULTS

In this section, we show the results of the differential evolution Markov chain and discrete cosine transform (DCT-DEMC) inversion algorithm applied to synthetic and real data. The forward modelling used was extracted from the boundless electrical resistivity tomography open-source software (Günther *et al*., 2006) and it is a 2.5D Finite Elements code implemented in Python.

Synthetic data inversion

To test the DCT-DEMC algorithm, we created a simple resistivity model composed of a low-resistivity (50 $\Omega \cdot$ m) rectangular block with the dimensions of 13 m $\times$ 2.5 m located at a depth of 1 m hosted in a higher resistivity medium (500 $\Omega \cdot$ m; see later in Fig. 5a). We simulate an acquisition using 36 electrodes 1 m spaced with the Wenner configuration; this choice agrees with the configuration commonly used in some long-term monitoring systems (e.g. Hojat *et al*., 2019; Tresoldi *et al*., 2019) where the Wenner array is used to guarantee high signal-to-noise ratio (Dahlin and Zhou, 2004). The dimension of the model is 35 m $\times$ 5.5 m composed of 1 m $\times$ 0.5 m cells; the total number of parameters in the uncompressed, resistivity domain is 385, whereas 11 data levels with a total number of 185 apparent resistivity data points were calculated for this model.

We contaminated the observed synthetic data with Gaussian random noise choosing a standard deviation that corresponds to 10% of the standard deviation of observed synthetic data. An important factor of the MCMC inversion is to build robust prior distribution based on previous knowledge about subsurface; for simplicity, we assume a Gaussian prior distribution calculating its moments (mean and standard deviation) directly from the synthetic model. For example, we calculate the mean by performing the arithmetic mean of synthetic model resistivity values in the logarithmic domain. The values of mean and standard deviation are 5.8259 and 0.8637, respectively (Table 1). The calculated moments are used with the covariance matrix (Equation 7) to generate prior model realizations. At this point, we need to set the optimal DCT coefficients number reducing the dimension of model space without losing spatial variability. For that reason, we draw six prior model realizations from the prior distribution, and for each model we investigate the spatial variability of the reconstructed model changing the retained number of DCT coefficients (Fig. 3).

We observe that the variability of 98% of the considered prior model realization is expressed considering five discrete cosine transform (DCT) coefficients along the k_x direction and three DCT coefficients along the k_y direction. Therefore, we retain the first five columns and the first three rows of the DCT coefficient matrix for a total of 15 coefficients setting $q = 3$ and $p = 5$ in $\mathbf{R}_{qp}$ matrix of equation (3). The compression reduces the 385-D original resistivity domain, to a 15-D space.

We run 10,000 iterations of the DCT-DEMC algorithm using four chains (Table 1). Besides, to illustrate the benefit of DCT reparameterization we compare the predictions provided by the DCT-DEMC algorithm with those yielded by a differential evolution Markov chain (DEMC) sampling running in the unreduced model space. In both cases, the starting model for each chain is drawn from the prior distribution.

Figure 4(a) shows the negative log-likelihood evolution for the DEMC inversions running in the compressed and uncompressed space. As expected, the model compression guarantees faster convergence towards the stationary regime that for the DCT-DEMC is attained in only 1000 iterations, while the DEMC running in the uncompressed model does not reach a data misfit values comparable to the DCT-DEMC within the selected number of iterations. This difference is related to the curse of dimensionality issue that is effectively mitigated by the DCT compression.

Figure 4(a) also illustrates that the models sampled in the first 1000 iterations must be neglected when computing the PPD to properly burn-in.

To quantitatively assess the convergence of the algorithms, we compare the potential-scale-reduction-factor (PSRF) computed for three subsurface resistivity cells (Fig. 4b). We note that less than 500 iterations are needed by the

Table 1 Parameters of the DCT–DEMC algorithm. The mean and the standard deviation, which are expressed in the logarithmic domain, are calculated from the synthetic resistivity model.a_x and a_y indicate the ranges of Gaussian variogram along x and y directions

Mean	Standard Deviation	Chains	Iterations	DCT Coefficients	a_x	a_y
5.8259	0.8637	4	10,000	15	2 m	2 m

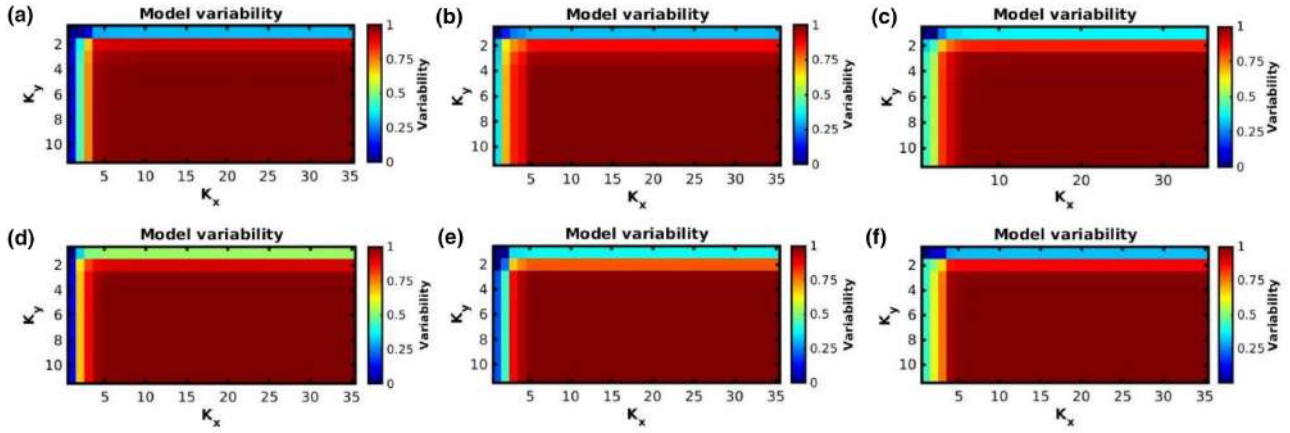

Figure 3 (a)–(f) Computed variability of six prior model realizations as the number of retained DCT coefficients along the k_y and k_x direction changes. Each model variability cell with coordinate (k_x, k_y) represents the variability of the reconstructed model computed considering the first k_x columns and k_y rows in the DCT coefficients matrix of equation (3).

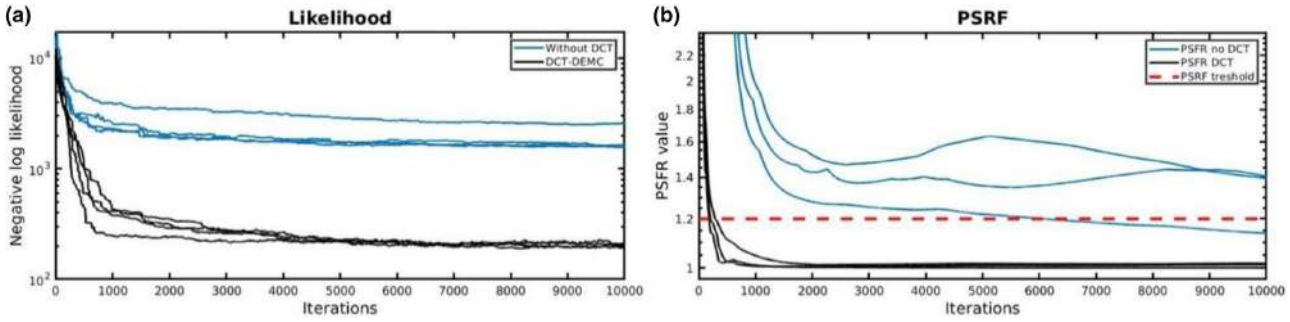

Figure 4 (a) Negative log-likelihood evolution for four chains and the DCT-DEMC (black) and DEMC (blue). (b) PSRF of three resistivity cells for the DCT-DEMC (black) and DEMC (blue) algorithm. The dotted line denotes the PSRF threshold of convergence.

DCT-DEMC to attain the PSRF threshold value of 1.2 that indicates convergence (Gelman *et al.*, 1995; Brooks *et al.*, 1998), whereas the standard DEMC algorithm does not reach that threshold within the selected number of iterations. Figure 5 compares the true model with the posterior mean models estimated by the two DEMC inversions. The DCT-DEMC provides a posterior mean in which the resistivity anomaly of the true model is correctly recovered. As expected, the quality of the results decreases towards the boundaries of the model due to poor illumination. The lower spatial resolution of the estimated model with respect to the true model is related to the resolution limits of ERT data, to the model compression technique, and also to the spatial variogram model infused into the prior assumptions. Significant scattering affects the posterior mean estimated by the DEMC running in the unreduced space (Fig. 5c), thus proving that this method fails to converge towards a stable posterior probability density (PPD) and that a model compression is needed to attain reliable results.

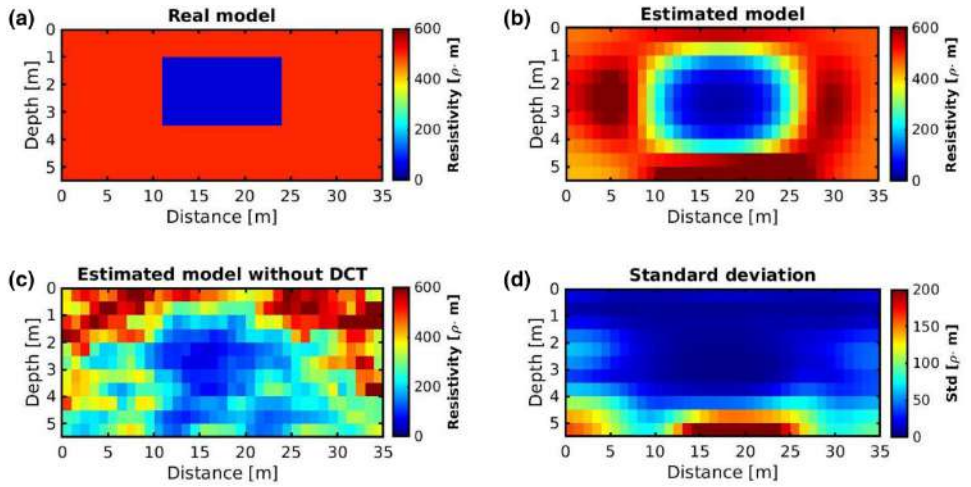

Figure 5 (a) Synthetic model. (b) DCT-DEMC estimated model. (c) Estimated model without DCT reparameterization. (d) Standard deviation associated with the DCT-DEMC results.

The analysis of the posterior standard deviation associated with the DCT-DEMC inversion (Fig. 5d) indicates an increase of uncertainties beyond 4 m depth at the lateral edges of the model, in agreement with the expected subsurface illumination. In other terms, the posterior standard deviation indicates that the cells located at lateral edges and below 4 m are not informed by the data and the associated resistivity values cannot be recovered with reasonable accuracy. The lowest values of the standard deviation are located within the central anomaly and in the shallowest part of the model (Fig. 5d). Figure 6 shows examples of posterior marginal distribution for six cells in the resistivity domain.

Figure 7 presents some examples of the resistivity model drawn from the estimated PPD. In all cases, we observe that the central low-resistivity anomaly is well recovered as well as the resistivity values in the shallowest part of the model. In addition, the difference between the posterior realization increases at the lateral edge and the bottom of the model. The posterior realizations can aid the interpretation phase because they fully capture the uncertainty in the estimated resistivity model or, in other terms, they represent possible subsurface scenarios in accordance with the acquired data.

Field data inversion

The increasing need to develop methods capable of giving early warning alarms for flood risks has resulted in the development of permanent monitoring systems that assess the structural health conditions of levees. The electrical resistivity tomography (ERT) technique, used in permanent monitoring systems, is appropriate to identify subsurface saturation changes, compositional changes and weak zones (Tresoldi *et al.*, 2019).

The field data that we invert in this paper was acquired along a critical section of the river levee in Colorno, Italy. The acquisition profile is composed of 48 plate electrodes using the Wenner array with the spacing of 2 m. The number of apparent resistivity points is 360, whereas the subsurface is discretized with rectangular cells 2 m long and 1 m thick with a maximum depth of 15 m; in this way, the total number of unknowns is 705. Preliminary knowledge and available information about the study area (Hojat *et al.*, 2019) helped us in setting the prior information; the process to build the prior realizations is schematized in Figure 2. We again assume a stationary log-Gaussian prior, whereas a Gaussian variogram is used to impose the lateral continuity to the prior realizations. The available previous information suggests the presence of the main clay body hosting gravels at shallow depth (above 4 m in depth). The data have a high signal-to-noise ratio; consequently, we perform only a mobile average filtering which attenuates the high-frequency content of pseudosection profiles.

As in the synthetic case, the optimal number of DCT coefficients to retain is evaluated exploring the variability of some resistivity models drawn from the prior distribution.

According to the model variability maps of Figure 8, if we consider 15 DCT coefficients along k_x and 10 DCT coefficients along k_y, this explains about 98% of the variability of the prior realizations. Therefore, the model space dimension is reduced from 705 to 150.

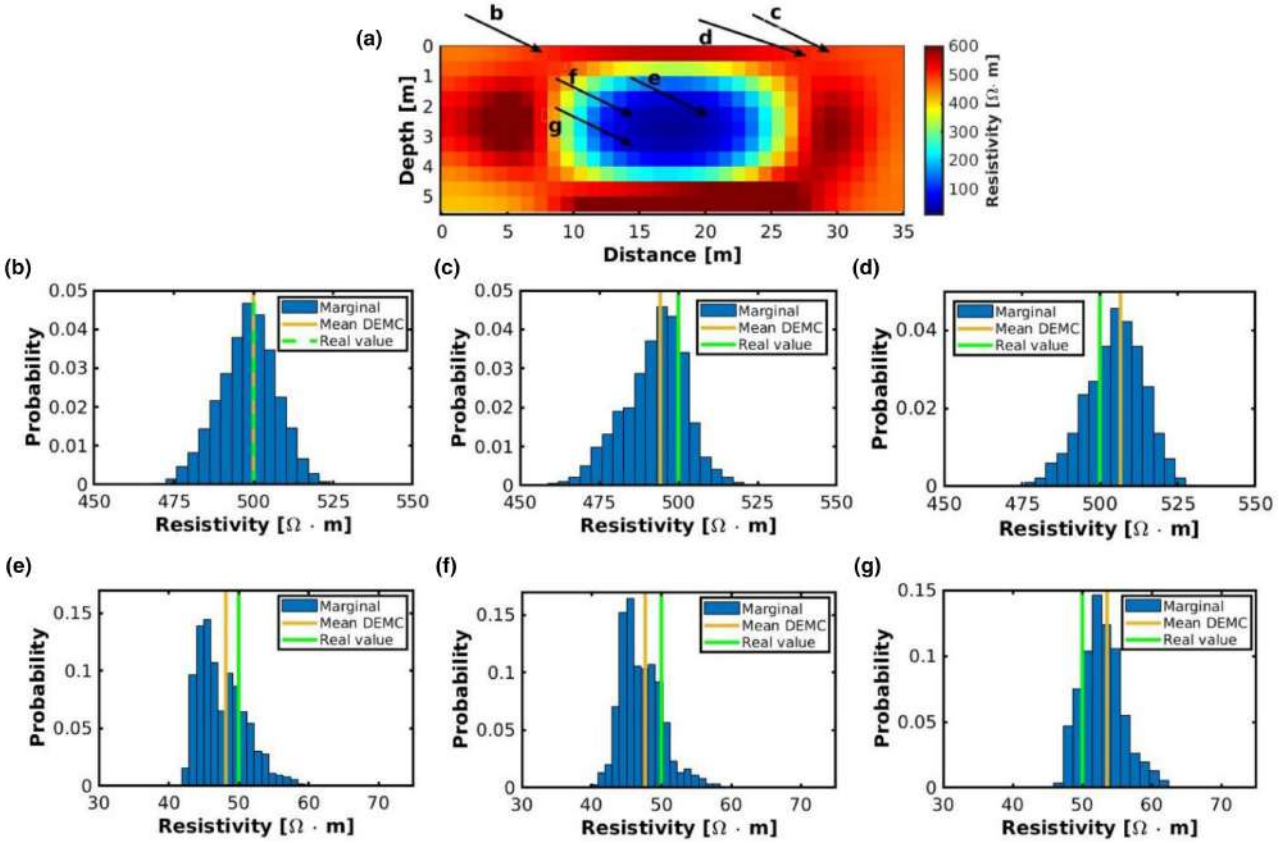

Figure 6 (a) Estimated model. The arrows indicate the selected resistivity cells. (b)–(d) Three marginal distributions associated with depth 0.5 m and distance 8, 30, and 28 m respectively. (e)–(f) marginal distributions associated with depth 2.5 m and distance 21 and 15 m, respectively. (g) Marginal distribution associated with the resistivity cell at depth 3.5 m and distance 15 m. The blue histograms are the marginal PPD function, the yellow vertical bar indicates the mean value of the marginal distribution and the green vertical bar represents the corresponding resistivity value of the synthetic model.

Table 2 Parameters of the DCT–DEMC algorithm. The mean and standard deviation, which are expressed in the logarithmic domain, are calculated from a model inverted with the gradient-based algorithm. a_x and a_y indicate the ranges of Gaussian variogram along x and y diections

Mean	Standard Deviation	Chains	Iteration	DCT Coefficients	a_x	a_y
3.4585	0.4223	20	50,000	150	3 m	2 m

We run the DCT-DEMC for 50,000 iterations and employing 20 independent chains that start from a model randomly drawn from the prior. A resume of the inversion parameters is given in Table 2. By observing the evolution of the negative log-likelihood shown in Figure 9(a), the first 15,000 iterations can be considered the burn-in period, while the successively sampled models have been used to numerically compute the posterior uncertainties. Again, the PSRF is used to assess the convergence of the sampling towards a stable PPD. Figure 9(b) shows examples of PSRF evolution for 10 model parameters in the DCT space, and it emerges that 40,000 iterations are needed to attain convergence.

To validate the DCT-DEMC outcomes, we compare the mean model estimated by the implemented approach with the prediction of a standard gradient-based inversion that implements the Levenberg–Marquardt (LM) method. For the LM inversion, we choose a starting model with constant resistivity calculated from the arithmetic mean of observed data. The similarity of the results provided by the deterministic and probabilistic approach demonstrates the applicability

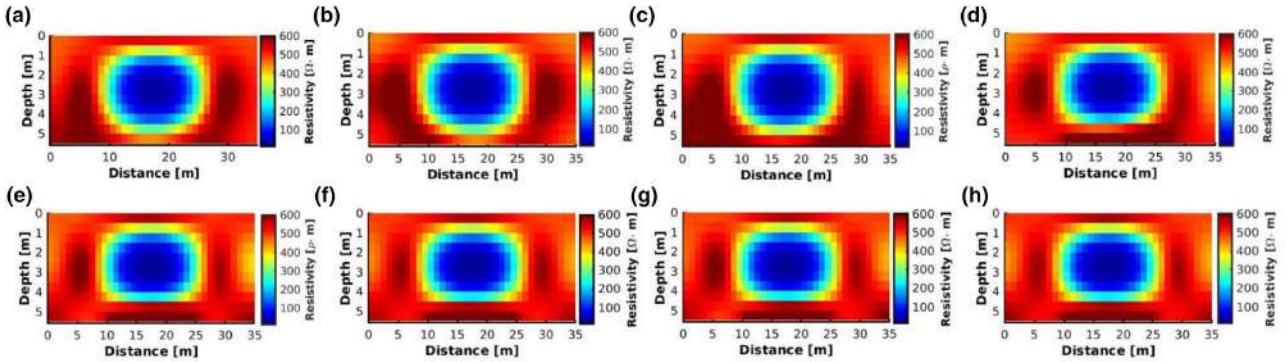

Figure 7 (a)–(h) Modelsdrawn from the posterior distribution (PPD).

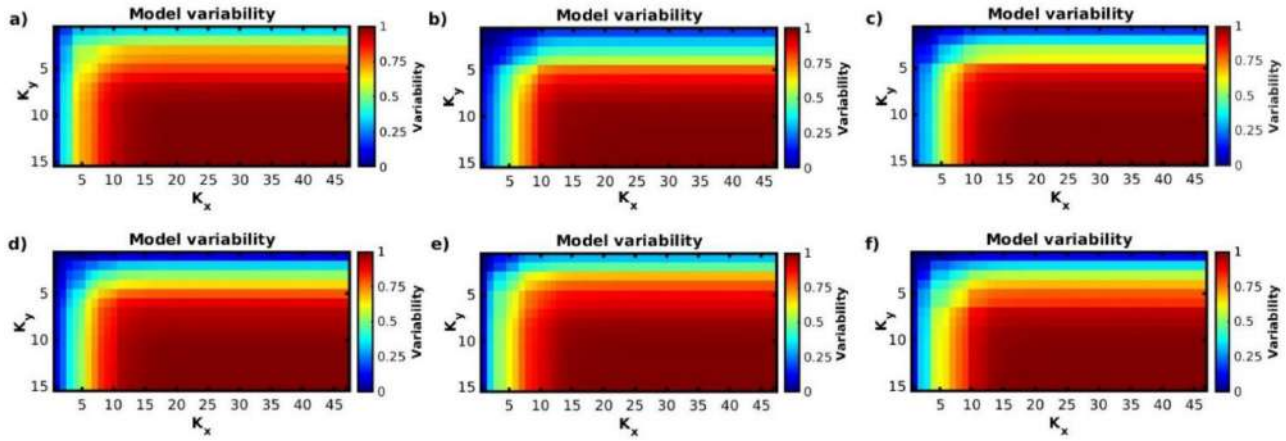

Figure 8 (a)–(f) Computed variability of six prior model realizations as the number of DCT coefficients along the k_y and k_x directions changes. Each cell with coordinates (k_x, k_y) represents the numerical value of reconstructed model variability computed retaining the first k_x columns and the first k_y rows of DCT coefficients.

of the proposed approach (Fig. 10a, b), while the major differences are localized at the lateral and bottom part of the model where the data illumination is poor. 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 LM approach, the DCT-DEMC inversion also provides hints about the uncertainty affecting the recovered model. The posterior standard deviation is small in the shallowest part and increases moving towards the deeper and lateral parts of the model, according to the expected sensitivity of the data to the subsurface resistivity values. Figure 11 presents eight statistical realizations drawn from the estimated posterior model. As for the synthetic inversion test, these realizations fully capture the uncertainty in the estimated resistivity model shown in Figure 10(b) and represent possible subsurface scenarios. Note that all the posterior realizations univocally predict a high-resistivity medium above 4 m depth.

Figure 12 shows a comparison between the observed data and the apparent resistivity sections computed on the mean model estimated by the DCT-DEMC and on the final solution of the LM approach. The two algorithms provide similar data matching although the data generated by the DEMC solution seems to be characterized by a slightly higher mismatch (rms = 5.2%) with respect to the observed pseudosection. This increased data misfit is related to the compression technique used to reduce the number of unknown parameters and to the consequent loss of model and data resolution.

To further investigate this higher misfit, we plot the observed data with the corresponding error bar (Figure 13) and the predicted data by LM and by the DCT-DEMC algorithm. The standard deviation has been extracted from the data

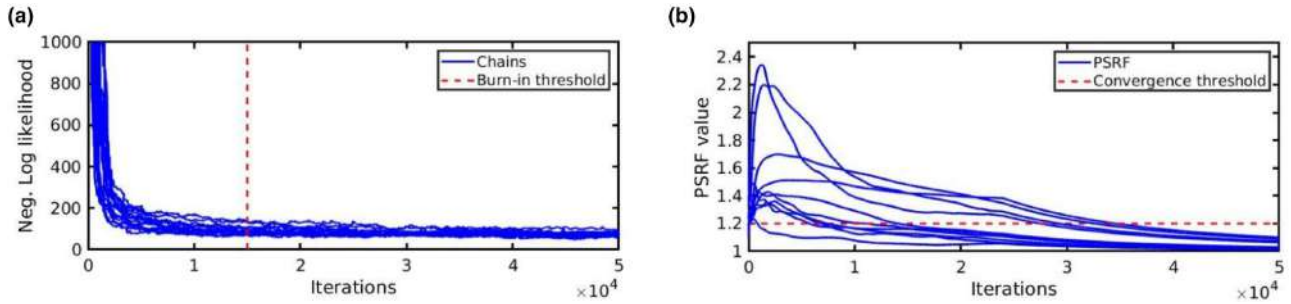

Figure 9 (a) Evolution of the negative log-likelihood for the 20 chains used in the real data case. The vertical red dashed line indicates the end of the burn-in period. (b) Evolution of the PSRF for 10 DCT coefficients. The horizontal red dashed line represents the threshold of convergence.

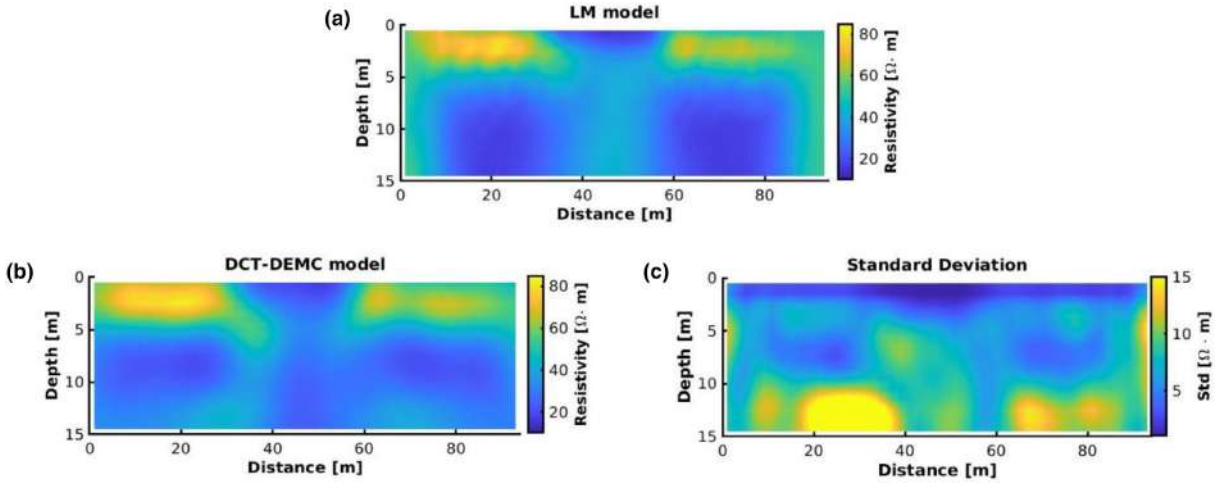

Figure 10 (a) Estimated model by the Levenberg Marquardt algorithm. (b) Posterior mean model provided by the DCT–DEMC algorithm. (c) Standard deviation estimated by DCT–DEMC.

covariance matrix employed for data likelihood evaluation (Equation 9). From this analysis emerges that both data associated with the gradient-based solution and posterior DCT-DEMC mean lie within the uncertainty range as depicted by standard deviation bars, thus proving that both methods reproduce the observed data equally well.

In Figure 14, we display examples of marginal PPD functions for six DCT coefficients. For the low-order coefficients, the posterior is different from the prior (calculated from Equation 5), thereby indicating that the low-frequency resistivity variation is well constrained by the data (Fig. 14a–c). On the contrary, the high-order coefficients associated with high-frequency resistivity variations are less resolved and this results in marginal posterior very similar to the prior (Figure 14d–f).

Figures 15 presents some examples of marginal prior and posterior distributions after projection in the resistivity domain. As expected, the model uncertainties increase moving from the shallowest (Fig. 15b–d) to the deeper part of the model (Fig. 15e–g). The good agreement between the LM predictions and the posterior mean estimated by the DCT-DEMC inversion again confirms the reliability of the proposed inversion approach.

Finally, we run two DCT-DEMC inversion tests: in the former, we reduce the number of retained coefficients to 100 ($q = 10$, and $p = 10$ in Equation 4), while the latter considers 200 coefficients in the compressed domain ($q = 10$, and $p = 20$) to explore the improvement of the spatial resolution in that direction. In both cases, we employ the same number of iterations (50,000), and the number of chains (20) of the previous inversion test. The inversions running with 150 and 200 unknowns attain similar negative log-likelihood values, while the 100-coefficient inversion results in a decreased data matching (Figure 16a).

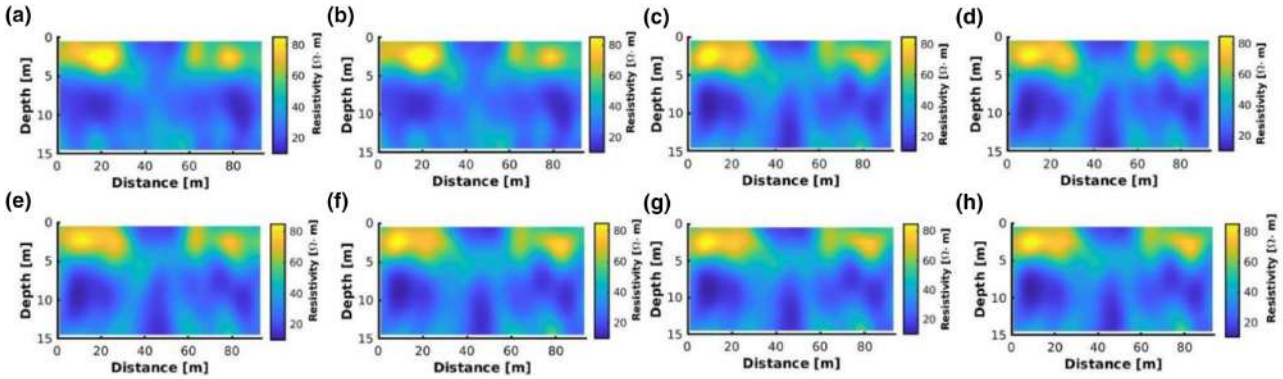

Figure 11 (a)–(h) Models drawn from the posterior distributions (PPD) estimated by the DCT–DEMC inversion.

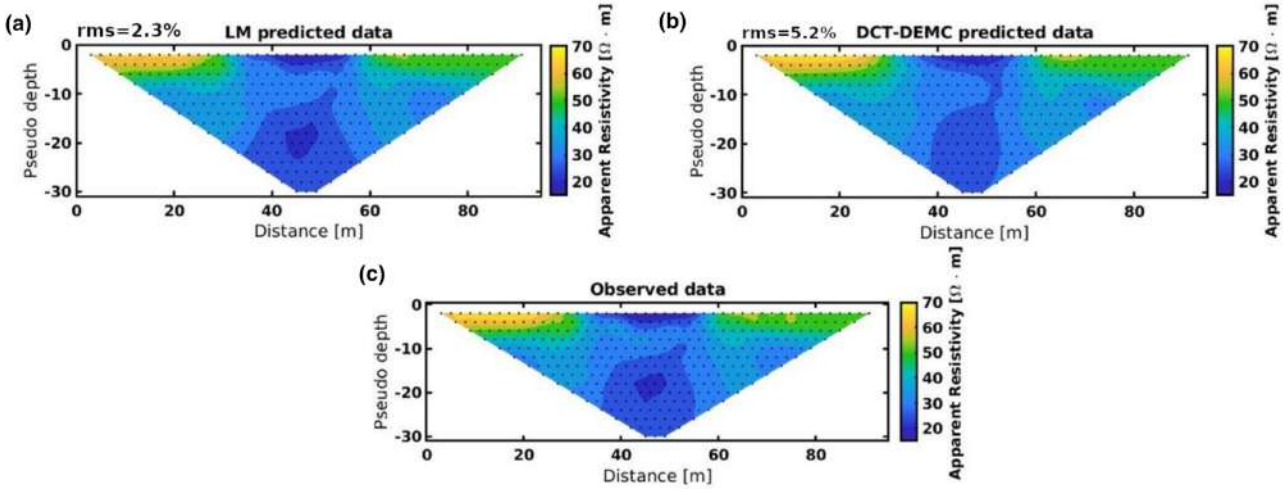

Figure 12 (a) Predicted data by the gradient-based algorithm (LM). (b) Predicted data by the DCT–DEMC algorithm inverting DCT coefficients. (c) Observed data.

The three inversions provide comparable posterior mean models (Fig. 16b–d) in which the shallow anomalies are well recovered, and, as expected, the spatial resolution increases as the number of retained coefficients increases. However, it can be noticed that the mean model estimated by the 200-coefficient inversion is very similar to the model estimated when 150 coefficients are employed. The predicted data computed on the mean models estimated by the three inversions are represented in Figures 16(e–g). Again, we note that the 100-coefficient inversion results in a higher data misfit value (rms data error of 6.2%) than those attained by the other two inversions (rms values of about 5%). For comparison, the data misfit for the LM model is 2.3%.

Figure 17 shows an example of PSRF evolution for different model parameters associated with the 150- and 200-coefficient inversion. This demonstrates that the number of iterations needed to attain stable PPD estimations significantly increases as the dimension of the model parameter increases. Therefore, this analysis together with the comparison shown in Figure 16 confirms that the inversion carried out with 150 coefficients constitutes the optimal compromise between model resolution, data fitting and the computational cost of the sampling procedure.

DISCUSSION

This work was aimed at casting the electrical resistivity tomography (ERT) inversion into a solid probabilistic framework for accurate uncertainty assessments. However, in ERT inversion the number of parameters is large, the problem is

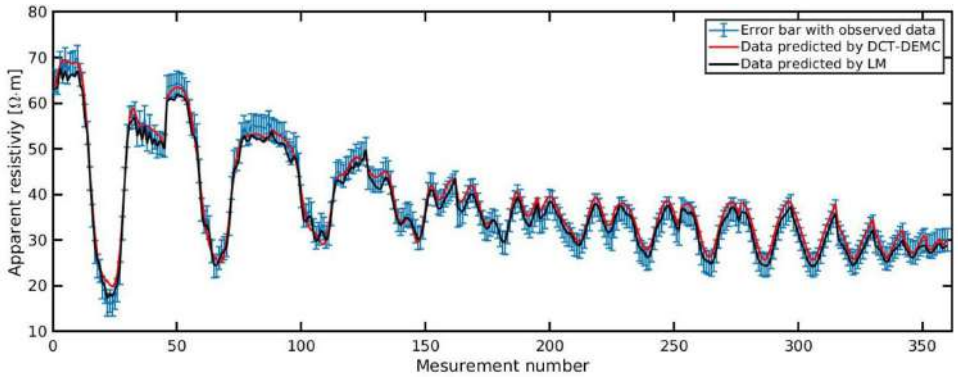

Figure 13 Standard deviation value associated with each observed data point (blue bar); predicted data by LM and the DCT–DEMC inversion algorithm (black and red curves, respectively).

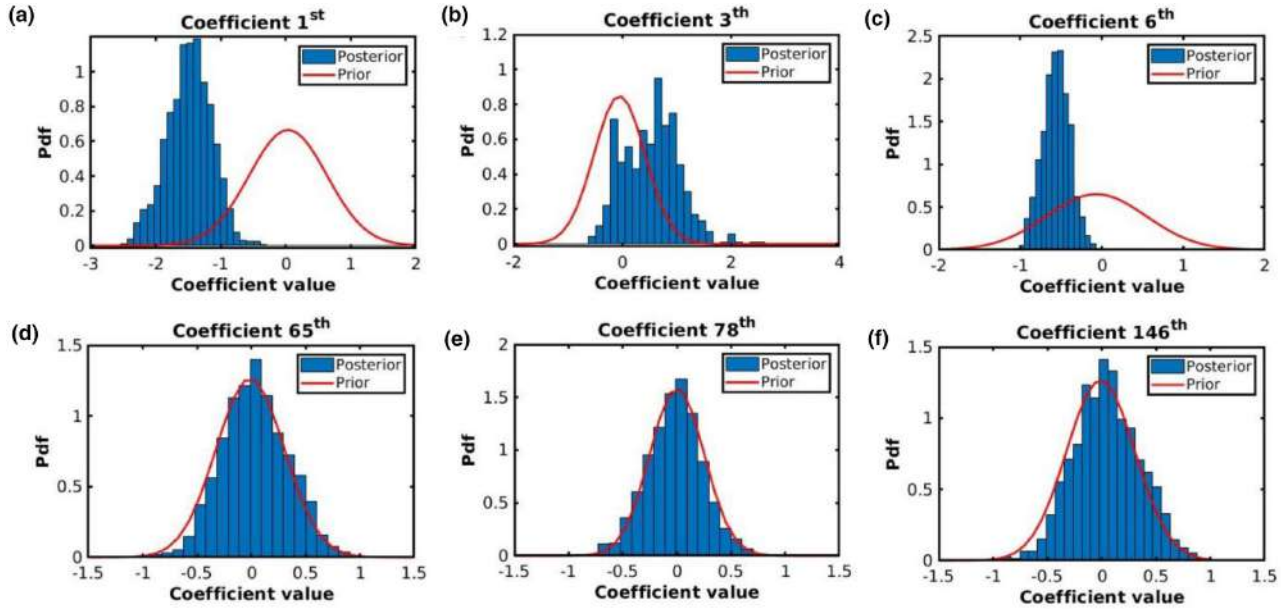

Figure 14 (a)–(c) Marginal PPDs for six DCT coefficients (blue bars) and associated prior distribution (red). Note where the prior distributions are almost coincident with the posterior distributions the coefficients are associated with a non-informative likelihood.

non-linear, and, consequently, the sampling of the posterior probability density (PPD) function is computationally prohibitive. For this reason, we combined an improved Markov chain Monte Carlo algorithm (the differential evolution Markov chain [DEMC]) with a discrete cosine transform (DCT) compression of the model space.

The choice of the number of DCT coefficients to compress the model space should always constitute a compromise between the desired spatial resolution and the dimensionality reduction of the parameter space. However, such a threshold level can be conveniently set by analysing how the number of retained coefficients is able to reconstruct the realization drown from the prior distribution. Imposing the error in prior model reconstruction to 98% allows in our examples to obtain synthetic data computed from the inverted model within the error bounds of the observed (synthetic or real) data.

The parallel tempering strategy (Dosso *et al.*, 2012) could be used to improve the mixing of the different DEMC chains, and this approach has proven to be particularly useful when sampling multimodal posteriors with modes separated by low probability regions. However, our experiments showed unimodal marginal posterior and for this reason we decided

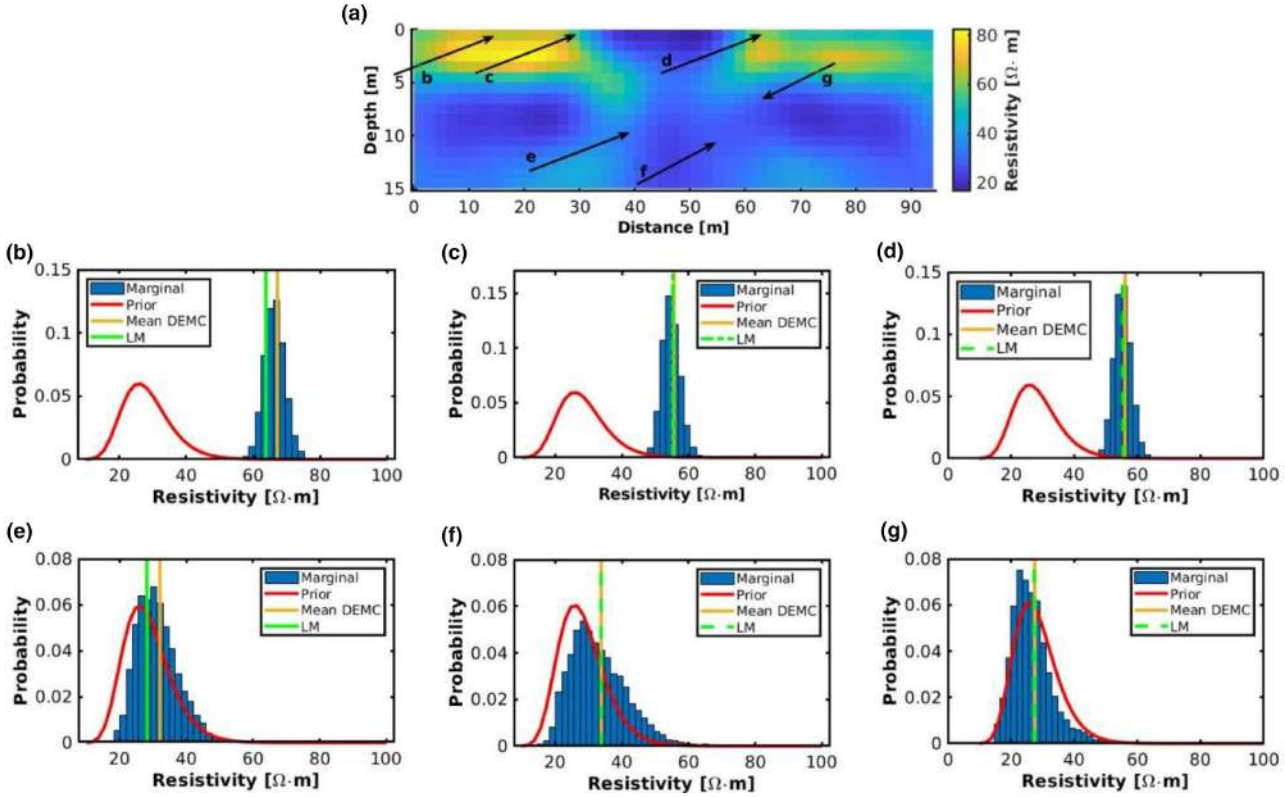

Figure 15 (a) Estimated model. The black arrows indicate the selected resistivity cells. (b)–(d) Marginal PPD of selected cells at depth equal to 1 m and distance 24, 30 and 62 m, respectively. (d)–(f) Marginal PPD of selected cells calculated at depths 10,11 and 7 m and distance equal to 40, 58 and 62 m, respectively. In each plot, the red curves represent the marginal prior probability density function, the vertical yellow lines represent the posterior mean model estimated by the DCT–DEMC algorithm, whereas the vertical green lines depict the LM solution.

not to include the parallel tempering within our sampling approach.

In this paper, to simplify the description we limit our attention to analytical prior, but another considerable benefit of the implemented approach is the possibility to derive accurate uncertainty appraisal for whatever type of prior assumption is desired (i.e. either parametric or non-parametric distributions).

The DCT is defined for multidimensional signals and for this reason the DCT-DEMC inversion can be easily extended to three-dimensional (3D) models. Obviously, in this case, fast forward modelling routines are needed to make the probabilistic sampling computationally affordable.

Indeed, the main computational requirement of the proposed approach lies in the need to run several forward model evaluations until a stable PPD is reached. In terms of computational cost, the DCT-DEMC code runs in 57 min for the synthetic case and in about 35 h for the real data inversion, considering a Python implementation running on a computer equipped with Intel i7-9700 CPU at 3.60 GHz 32 Gb RAM.

The boundless electrical resistivity tomography forward modelling code uses the CHOLMOD package (Davis, 2006) that aims to minimize the computer storage required by the sparse Cholesky method (George and Liu, 1981) to calculate the potentials. The CHOLMOD code is very efficient for large meshes such as for 3D finite-element models. However, it uses a more complex data structure than the simpler band and envelope methods (George and Liu, 1981) that introduces some computational overhead. The two-dimensional finite-element mesh used for surveys with electrodes on the ground surface usually involves smaller meshes where the number of nodes in the horizontal direction is much greater than in the vertical direction. Initial research indicates that code (using the simpler sparse Cholesky methods) optimized for such 'long and thin' meshes can be an order of magnitude faster.

As an alternative, the DCT-DEMC inversion can be started from a preliminary estimate of the subsurface

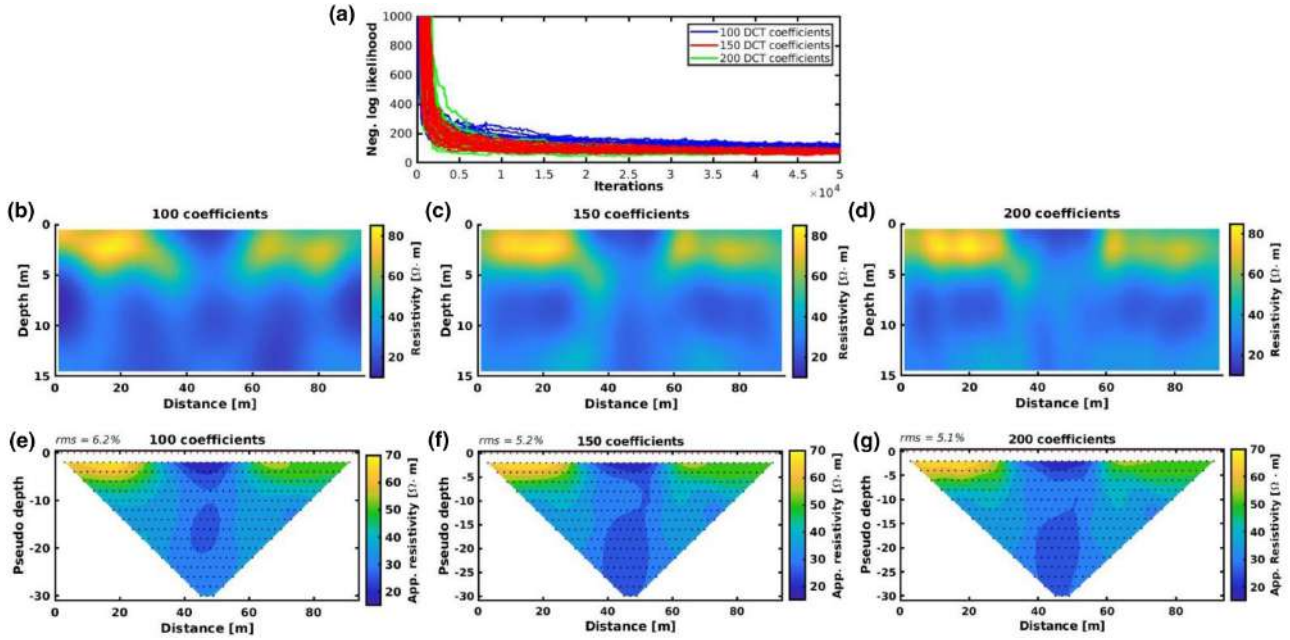

Figure 16 (a) Negative log-likelihood evolution for 100 DCT coefficients inversion (blue), 150 DCT coefficients inversion (red) and 200 DCT coefficients inversion (light green). (b)–(d) Posterior mean model estimated by the 100-,150- and 200-coefficient inversion, respectively. (e)–(g) apparent resistivity pseudosections generated by the model represented in (b)–(d), respectively.

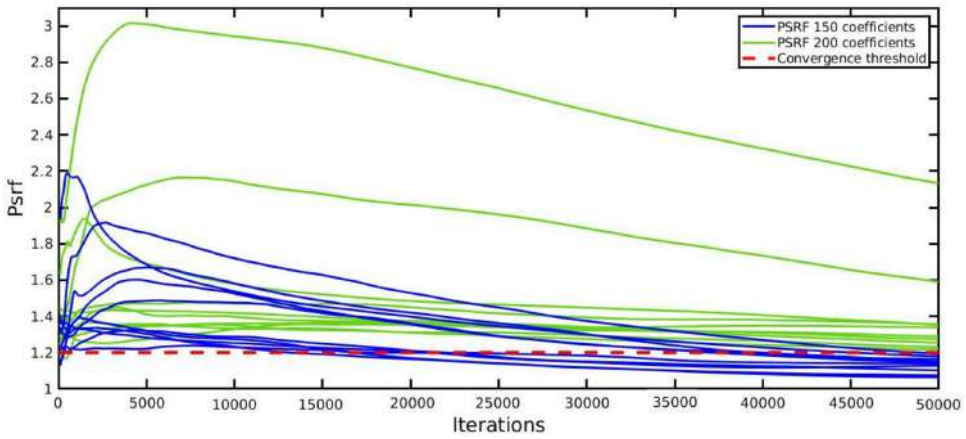

Figure 17 Examples of PSRF evolution for 10 DCT coefficients pertaining to the 150- and 200-coefficient inversions (blue and green curves, respectively). The red dashed line indicates the threshold of convergence.

resistivity models as provided, for example, by a standard gradient–based approach. This significantly reduces the length of the burn-in phase, thereby reducing the computational cost of the sampling procedure.

The proposed inversion approach is not aimed at replacing standard gradient–based inversion algorithms, but it could be used to attain accurate uncertainty estimations with a reasonable computational workload. The estimated PPD can aid the interpretation phase because gives a hint on the uncertainties affecting the recovered subsurface resistivity model.

CONCLUSIONS

In this paper, we proposed an alternative approach to the classical gradient-based electrical resistivity tomography inversion, that makes use of the Markov chain Monte Carlo probabilistic framework embedded in the differential

evolution Markov chain sampling. The main advantage of this procedure consists of the description of the solution as a posterior probability density function that allows the uncertainty appraisal of the estimated model parameters. To reduce the computational burden, the model is re-parameterized in terms of the coefficients of the discrete cosine transform (DCT), thereby reducing the number of unknowns to a level that could be easily managed computationally. The DCT reparameterization also acts as a model regularization strategy that preserves reasonable spatial continuity in the recovered solution.

Synthetic and field inversions demonstrated that the implemented approach provides reliable predictions in agreement with the outcomes of a gradient-based inversion. However, the outstanding benefit of the differential evolution Markov chain and discrete cosine transform algorithm is the possibility to accurately assess the uncertainties affecting the recovered solution. We deem that the possibility to estimate model uncertainties is worth the additional computational effort required.

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. 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. Open Access Funding provided by Universita degli Studi di Firenze within the CRUI-CARE Agreement.

DATA AVAILABILITY STATEMENT

Data are available on request from the authors.

REFERENCES

Ahmed, N., Natarajan, T. and Rao, k. (1974) Discrete cosine transform. *IEEE Transactions on Computers*, C-23(1), 90–93. https://doi.org/10.1109/T-C.1974.223784

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. (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. and 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. and Pierini, S. (2020) Transdimensional and Hamiltonian Monte Carlo inversions of Rayleigh-wave dispersion curves: a comparison on synthetic datasets. *Near Surface Geophysics*, 18(5), 515–543.

Aleardi, M. (2020a) Combining discrete cosine transform and convolutional neural networks to speed up the Hamiltonian Monte Carlo inversion of pre-stack seismic data. *Geophysical Prospecting*, 68, 2738–2761.

Aleardi, M. (2020b) Discrete cosine transform for parameter space reduction in linear and non-linear AVA inversions. *Journal of Applied Geophysics*, 179, 104106.

Azevedo, L., Tompkins, M.J. and Mukerji, T. (2016) Regularized sparse-grid geometric sampling for uncertainty analysis in nonlinear inverse problems. *Geophysical Prospecting*, 64(2), 320–334.

Brooks, S.P. and Gelman, A. (1998) General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4), 434–455

Craiu, R.V. and Meng, X.L. (2005) Multi-process parallel antithetic coupling for forward and backward Markov chain Monte Carlo. *The Annals of Statistics*, 33, 661–697.

Curtis, A. and Lomax, A. (2001) Prior information, sampling distribution, and curse of dimensionality. *Geophysics*, 66(2), 372–378.

Dahlin, T. and Zhou, B. (2004) A numerical comparison of 2D resistivity imaging with 10 electrode arrays. *Geophysical Prospecting*, 52, 379–398.

Davis, T.A. (2006) *Direct Methods for Sparse Linear Systems*. Society for Industrial and Applied Mathematics.

Dosso, S.E., Holland, C.W. and Sambridge, M. (2012) Parallel tempering for strongly nonlinear geoacustic inversion. *The Journal of the Acoustical Society of America*, 132(5), 3030–3040.

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.

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.

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

George, A. and Liu, J.W., (1981) *Computer Solution of Large Sparse Positive Definite Systems*. Prentice-Hall.

Goldie, M. (2002) Self-potentials associated with the Yanacocha high-sulfidation gold deposit in Peru. *Geophysics*, 67, 684–689.

Grana, D., Passos de Figueiredo, L. and Azevedo, L. (2019) Uncertainty quantification in Bayesian inverse problems with model and data dimension reduction. *Geophysics*, 84(6), M15-M24.

Günther, T., Rücker, C. and Spitzer, K. (2006) Three-dimensional modelling and inversion of dc resistivity data incorporating topog-

raphy – II. Inversion. *Geophysical Journal International*, 166, 506–517.

Hastings, W.K. (1970) Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1), 97–109.

Helene, L.P.I., Moreira, C.A. and Bovi, R.C. (2020) Identification of leachate infiltration and its flow pathway in landfill by means of electrical resistivity tomography (ERT). *Environmental Monitoring and Assessment*, 192, 249. https://doi.org/10.1007/s10661-020-8206-5.

Hojat, A., Arosio, D., Longoni, L., Papini, M., Tresoldi, G. and Zanzi, L. (2019) 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. EAGE, Vol. 2019, pp. 1–5. https://doi.org/10.3997/2214-4609.201900421.

Holmes, C., Krzysztof, L. and Pompe, E. (2017) Adaptive MCMC for multimodal distributions.

Hübner, R., Heller, k., Günther, T. and Kleber, A. (2015) Monitoring hillslope moisture dynamics with surface ERT for enhancing spatial significance of hydrometric point measurements. *Hydrology and Earth System Sciences*, 19, 225–240

Jain, A.K. (1989) *Fundamentals of Digital Image Processing*. Prentice Hall.

Loke, M.H., Chambers, J.E., Rucker, D.F., Kuras, O. and Wilkinson, P.B. (2013) Recent developments in the direct-current geoelectrical imaging method. *Journal of Applied Geophysics*, 95, 135–156.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H. and Teller, E. (1953) Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1087–1092.

Mogadhas, D. and Vrugt, J.A. (2019) The influence of geostatistical prior modeling on the solution of DCT-based Bayesian inversion: a case study from Chicken Creek catchment. *Remote Sensing*, 11, 1549. https://doi.org/10.3390/rs11131549

Nunes, R., Azevedo, L. and Soares, A. (2019) Fast geostatistical seismic inversion coupling machine learning and Fourier decomposition. *Computational Geosciences*, 23(5), 1161–1172.

Pelton, W. and Smith, P. (1976) Mapping porphyry copper deposits in the Philippines with IP. *Geophysics*, 41, 106–122.

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

Ramirez, A.L., Nitao, J.J., Hanley, W.G., Aines, R., Glaser, R.E., Sengupta, S.K., et al. (2005) Stochastic inversion of electrical resistivity changes using a Markov chain Monte Carlo approach. *Journal of Geophysical Research: Solid Earth*, 110(B2), B02101.

Sambridge, M. and Mosegaard, K. (2000) Monte Carlo methods in geophysical inverse problems. *Reviews of Geophysics*, 40, 3.

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

Scalzo, R., Kohn, D., Olierook, H., Houseman, G., Chandra, R., Girolami, M., et al. (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.

Stuart, G.K., Minkoff, S.E. and Pereira, F. (2019) A two-stage Markov chain Monte Carlo method for seismic inversion and uncertainty quantification. *Geophysics*, 84(6), R1003–R1020.

Sumner, J. (1976) *Principles of Induced Polarization for Geophysical Exploration*. Elsevier Scientific.

Szabó, N.P. and Dobróka, M. (2019) Series expansion-based genetic inversion of wireline logging data. *Mathematical Geosciences*, 51(6), 811–835.

Ter Braak, C.A. (2006) Markov chain Monte Carlo version of the genetic algorithm differential evolution: easy Bayesian computing for real parameter spaces. *Statistics and Computing*, 16, 239–249.

Tarantola, A. (2005) *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.

Tresoldi, G., Arosio, D., Hojat, A., Longoni, L., Papini, M. and Zanzi, L. (2018) Tech-Levee-Watch: experimenting an integrated geophysical system for vulnerability assessment of levees. *Rendiconti online della Società Geologica Italiana*, 46, 38–43, https://doi.org/10.3301/ROL.2018.49.

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, https://doi.org/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.

Uhlemann, S., Kuras, O., Richards, L., Naden, E. and Polya, D. (2017) Electrical resistivity tomography determines the spatial distribution of clay layer thickness and aquifer vulnerability, Kandal Province, Cambodia. *Journal of Asian Earth Sciences*, 147, 402–414.

Vrugt, J. (2016) Markov chain Monte Carlo simulation using the DREAM software package: theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, 75. 273–316 https://doi.org/10.1016/j.envsoft.2015.08.013.

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, 357–382.

