Geophysical Journal International

Geophys. J. Int. (2024) **239,** 1576–1596
Advance Access publication 2024 September 26
GJI Applied and Marine Geophysics

https://doi.org/10.1093/gji/ggae347

Uncertainty quantification in electrical resistivity tomography inversion: hybridizing block-wise bootstrapping with geostatistics

Zahra Tafaghod Khabaz [1], Reza Ghanati [1] and Charles L. Bérubé [2]

[1] *Department of Earth Physics, Institute of Geophysics, University of Tehran, 14359-44411 Tehran, Iran. E-mail: rghanati@ut.ac.ir*
[2] *Polytechnique Montréal Civil, Geological and Mining Engineering Department, H3C 3A7 Montréal, Canada*

Accepted 2024 September 20. Received 2024 August 29; in original form 2024 June 11

SUMMARY

Electrical resistivity tomography inversion often encounters uncertainty stemming from two primary sources: epistemic uncertainty, arising from imperfect underlying physics and improper initial approximation of model parameters, and aleatory variability in observations due to measurement errors. Despite the widespread application of electrical resistivity tomography in imaging, the resistivity distribution of subsurface structures for various hydro-geophysical and engineering purposes, the assessment of uncertainty is seldom addressed within the inverted resistivity tomograms. To explore the combined impact of epistemic and aleatory uncertainty on resistivity models, we initially perturb the observed data using non-parametric block-wise bootstrap resampling with an optimal choice of the block size, generating different realizations of the field data. Subsequently, a geostatistical method is applied to stochastically generate a set of initial models for each bootstrapped data set from the previous step. Finally, we employ a globally convergent homotopic continuation method on each bootstrapped data set and initial model realization to explore the posterior resistivity models. Uncertainty information about the inversion results is provided through posterior statistical analysis. Our algorithm's simplicity enables easy integration with existing gradient-based inversion methods, requiring only minor modifications. We demonstrate the versatility of our approach through its application to various synthetic and real electrical resistivity tomography experiments. The results reveal that this approach for quantifying uncertainty is straightforward to implement and computationally efficient.

Key words: Electrical resistivity tomography; Inverse theory; Numerical modelling; Hydrogeophysics.

1 INTRODUCTION

Electrical resistivity tomography (ERT) is a non-invasive and near-surface geophysical method widely employed for spatially continuous imaging of the subsurface. It finds extensive application in various geoscientific endeavors, including the exploration of subsurface resources, environmental and engineering investigations and the imaging of hydrogeophysical properties. From both deterministic and probabilistic points of view, geophysical inversion aims to infer subsurface physical properties from limited and noisy data. Similar to numerous other nonlinear geophysical inverse problems, ERT data inversion frequently entails uncertainty (Tso *et al.* 2017). This uncertainty comprises a statistical component, termed aleatory uncertainty, which accounts for inherent noise or stochasticity in the observations, and a deterministic component, known as epistemic uncertainty, arising from incomplete knowledge about underlying physics (e.g. incorrect assumptions regarding isotropy, anisotropy, homogeneity and numerical approximation) and initial approximation of model parameters. Hence, to ensure reliable predictions of subsurface structures through the interpretation of electrical resistivity, it is imperative to quantify the uncertainty inherent in the inverted resistivity tomogram. In this study, our focus lies on uncertainties related to data errors and local variations in the initial model utilized for inversion. Despite the importance of uncertainty analysis, the majority of studies focusing on ERT data inversion have neglected a crucial factor, which is assessing the uncertainty in the preferred inversion model. Recalling the ERT inversion procedures, uncertainty quantification can be classified into two primary groups: deterministic linearized inversion and stochastic inverse algorithms. In deterministic linearized inversion, uncertainty analysis relies on a single solution that primarily reflects model uncertainty attributed to data noise, but does not consider inherent solution non-uniqueness. Stochastic inverse algorithms, on the other hand, aim to construct an ensemble of solutions. The latter provide a set of

models that sufficiently fit the observed data, resulting in statistical distributions of model parameters.

In regards to uncertainty estimation in ERT inversion, several studies have been reported in the literature. For instance, Ramirez *et al.* (1995) utilized the diagonal elements of the resolution matrix as proxies for the spatial resolution of ERT images in cross-well dc—direct current-resistivity measurements. Through the examination of these diagonal elements, they could qualitatively assess regions demonstrating satisfactory resolution. Friedel (2003) suggested the truncated singular value decomposition method to calculate model resolution and model covariance matrices. Using the resolution matrix, he computed metrics like the resolution radius and distortion indicator. These metrics were utilized to adjust the diagonal elements of the resolution matrix, thereby identifying possible geometric distortions in ERT tomograms. Slater *et al.* (2000) took advantage of the diagonal values of the model resolution matrix to distinguish poorly and perfectly resolved parts of the inverted conductivity model in a 2D resistivity cross-hole measurements. Oldenburg & Li (1999) devised a depth of investigation method, generating a series of distinct deterministic models from the same data set through repeated deterministic inversions with varying regularization constraints. Features consistently present across all models are interpreted as being effectively resolved by the data. Unlike the model resolution matrix, their method does not depend on the linearization assumptions. Later, Oldenborger *et al.* (2007) proposed an extension of the depth of investigation method to 3D electrical imaging problems, expressed as a volume of investigation. Oldenborger & Routh (2009) investigated the appraisal problem for 3D electrical resistivity measurements using analysis of the point spread functions of the linearized resolution matrix. The point spread functions represent the impulse response of the inverse solution and quantify the parameter-specific resolving capability. A modified depth of investigation method based on the computation of a scaled probability density function was developed by Deceuster *et al.* (2014). Kalscheuer *et al.* (2010) calculated the reliability of 2D resistivity models computed using smoothness-constrained single and joint inversions of ERT and radio magnetotelluric (RMT) data sets comparing the results of linearized analysis and most-squares inversion. They showed that equivalent model domains can be calculated using the partially nonlinear concept of pseudo-hyperellipsoids.

Another approach to ERT image uncertainty analysis involves the cumulative sensitivity matrix, which enables the examination of how surface or borehole measurements are affected by the parameters of the model cells (Kemna 2000). Typically, areas characterized by low-sensitivity values are regarded as less reliable. According to Kemna (2000), Christiansen & Auken (2012), Nguyen *et al.* (2009), and Beaujean *et al.* (2010), a significant challenge in utilizing the cumulative sensitivity matrix is establishing an objective and quantitative approach to selecting a threshold value that distinguishes between well and poorly resolved areas, especially in scenarios where ground truth data is unavailable.

An alternative strategy is to cast the problem in a stochastic framework using Bayesian inference (Mosegaard & Tarantola 1995; Gouveia & Scales 1997). Bayesian inference provides a formal framework for incorporating data uncertainties, forward models, and a priori information to quantify the uncertainty of model parameters by sampling the posterior probability density function of each model parameter using techniques such as Markov chain Monte Carlo (McMC) methods.

Within the context of ERT, several authors have utilized the McMC sampling algorithm to quantify the uncertainty of parameter estimates (e.g. Ramirez *et al.* 2005; Rosas-Carbajal *et al.* 2014; Bouchedda *et al.* 2017; Galetti & Curtis 2018; de Pasquale *et al.* 2019; Aleardi *et al.* 2021; Vinciguerra *et al.* 2021).

However, solving inverse problems in the framework of stochastic approaches is computationally demanding in practice, primarily due to the repeated computation of the forward problem, which must be solved tens or hundreds of thousands of times, and extremely slow convergence speed (see Blatter *et al.* (2022a) for more drawbacks of McMC). Although the enhanced computational capabilities provided by modern parallel architectures have significantly spurred the utilization of McMC methods in tackling geophysical challenges, it remains essential to employ tailored methodologies to ensure precise and computationally efficient sampling of the posterior probability distribution. In addition, Bayesian methods essentially necessitate an explicit statistical representation of the model error via the formulation of a likelihood function. However, this can pose a significant challenge for complex geophysical models, especially when there is limited or no information available to infer the prior distribution of the error model. Within the realm of multisolution methods, another approach involves generating realizations of the presumed noise model and subsequently perturbing the data set with these realizations within a Monte Carlo framework. Each perturbed data set is then inverted using a least-squares inversion method (Aster *et al.* 2005; Tso *et al.* 2017). Nevertheless, this method often yields conservative uncertainty estimates in the inverted model, particularly in regions of low-data sensitivity that are heavily influenced by the regularization operator (Binley & Slater 2020). As an alternative to McMC methods, ensemble Kalman inversion combined with level-set parametrization has been proposed for ERT and induced polarization data inversion and uncertainty quantification (Tso *et al.* 2021, 2024). However, despite its computational efficiency, it faces the common limitations of ensemble Kalman filter methods. For instance, using a limited number of samples can lead to the collapse of posterior samples, resulting in an underestimation of posterior uncertainty. Additionally, the assumption of Gaussian-distributed posteriors may break down in nonlinear scenarios. In the context of deterministic inversion limitations, as demonstrated by Wei & Sun (2021), various components of the regularization term influence the uncertainty quantification of physical property models. Therefore, a question that arises here is how to have a reliable uncertainty analysis in a relatively computationally efficient deterministic framework. In this paper, by admitting that uncertainty in geophysical models mainly stems from both the noise inherent within the observed data and the initial (prior) model, we introduce a hybrid approach that combines a variant of block bootstrapping, namely circular block bootstrap (CBB) (Politis & Romano 1992) with a geo-statistical method to address uncertainty in the regularized inversion of ERT data. The objective is to explore and sample the various solutions within the equivalence region, rather than opting solely for the simplest, best-fitting maximum likelihood solution. However, it is important to note that sampling the equivalence domain using the proposed method does not yield the desired Bayesian posterior distribution, which accounts for all significant sources of uncertainty, but rather serves as an approximation of it. Despite this limitation, we cannot definitively assert the superiority of one uncertainty quantification method over another. It is essential to recognize that uncertainty quantification is inherently local, not global and that prior assumptions, such as implicit or explicit regularization and model parametrization, will invariably have a significant impact on posterior uncertainty (Blatter *et al.* 2022a, b).

The bootstrap resampling algorithm has found widespread practical applications in various fields, notably in statistical inference (such as distribution functions, hypothesis tests, and model detection), as extensively documented in the mathematical literature. Despite its proven effectiveness, its utilization for quantitatively assessing uncertainty in geophysical models has been relatively rare. Nevertheless, some studies have explored this avenue (see McLaughlin 1988; Tichelaar & Ruff 1989; Shearer 1997; Hertrich 2008; Campanya *et al.* 2014; Neukirch & Garcia 2014; Yang *et al.* 2014; Parsekian-and-Grombacher 2015; Schnaidt & Heinson 2015). However, these studies primarily focus on addressing the uncertainty of geophysical models using conventional bootstrap resampling methods and, more importantly, none of them delve into the influence of epistemic uncertainty due to variability in the initial model.

The proposed stochastic strategy is executed in two stages. First, we employ CBB to construct an ensemble of randomly resampled data. Unlike other block bootstrap methods, CBB offers the advantage of wrapping data around in a circular before blocking, ensuring that each observation has an equal likelihood of being selected. While the efficacy of a block bootstrap method heavily relies on the selection of block length, we adopt the approach suggested by Politis & White (2004) to estimate the optimal block size. In the second stage, a geostatistical approach employing ordinary kriging is conducted on each bootstrapped data realization to statistically produce an initial guess model. Ultimately, each bootstrapped data set and its corresponding initial model are inverted by utilizing a globally convergent homotopic continuation method. The homotopic method, a longstanding element of topology, has been widely utilized in areas such as nonlinear functional analysis, differential geometry and engineering (Rion & Brunt 1999; Liao 2004; Axelsson *et al.* 2015). The primary benefit of the Homotopic method lies in its global convergence, ensuring convergence irrespective of the initial point chosen (Watson & Haftka 1989; Han *et al.* 2005). The research presented in this study is, to the best of the authors' knowledge, the first attempt to integrate the CBB method with Geostatistics to address the combined effects of epistemic and aleatory uncertainty in ERT inversion.

In the following sections, we provide a concise overview of the ERT forward modelling solution and the Homotopic continuation inversion method. This method is employed for the inversion of each data realization. Next, we delve into the fundamentals of the block bootstrap technique, with a particular focus on circle block bootstrapping. This is followed by an explanation of a geo-statistical approach used to stochastically generate a set of initial models for each bootstrapped data set. Then, we assess the performance of the proposed method through controlled synthetic experiments as well as practical applications using two distinct ERT data sets, illustrating the impacts of both data errors and initial models on uncertainty quantification in the tomographic inversion of electrical resistivity. Finally, concluding remarks are presented.

2 MATERIALS AND METHODS

In this section, we outline the mathematical and computational framework of how the proposed uncertainty quantification algorithm works. To attain our objective, we combine the regularized homotopic inversion with the proposed random resampling approach. For the remainder of this paper, we adopt the following conventions: we denote scalar quantities in italics, vectors in boldface lowercase letters, matrices in boldface uppercase letters, and the q-norm of a vector is represented by $\|.\|_q$.

2.1 Bootstrap resampling

Bootstrapping is a computationally intensive statistical method that offers a flexible alternative to traditional parametric inference by relaxing certain conditions and assumptions. As a subset of resampling techniques, bootstrapping enables the estimation of population statistical properties through sampling from an approximate or empirically constructed distribution, bypassing the necessity for prior assumptions. The idea of the standard bootstrap resampling was introduced by Efron (1979) as a strategy to estimate confidence intervals for model parameters. The traditional bootstrap method creates samples by randomly resampling the observed data, with or without replacement, under the assumption that the observations are independently and identically distributed, and then constructs the corresponding empirical distribution function.

Suppose $d \in \mathcal{R}^{m\times \mathbf{1}} = (d_1,\ d_2,\ \ldots,\ d_m)$ to be a sequence of independent and identically distributed random samples (observations) of size m from an unknown distribution. A bootstrap realization $\hat{d} = (\hat{d}_1,\ \hat{d}_2,\ \ldots,\ \hat{d}_m)$ is generated by randomly drawing p (where, $p < m$) samples from d without replacement, for each bootstrapped data set. To create multiple bootstrapped data sets, the resampling process is repeated until a total of k realizations are obtained $D = (\hat{d}_1,\ \hat{d}_2,\ \ldots,\ \hat{d}_k)$. Each of these realizations represents a subset of the original data set, formed by random sampling. Using partial data introduces ambiguity into the inversion results, enabling efficient sampling within the equivalent model domain (Fernández-Muñiz *et al.* 2019). The conventional bootstrap method assumes that data are independent and identically distributed. However, in some cases, correlations between the data can be observed, where the block bootstrap method partially maintains the underlying dependence structure and creates more realistic pseudo-samples. As a result, the performance of the original bootstrap method can be affected by dependence and correlation of observations. A remedy to this problem is to use the block-wise bootstrap to enhance the accuracy of bootstrap resampling by dividing the data samples into multiple blocks (subsets), where the blocks may be made up of non-overlapping or overlapping subsets from the original observed data (see Valavi *et al.* 2019) for additional advantages of splitting the data into subsets for random resampling).

In this paper, we follow the strategy of the overlapping blocks namely CBB which is an asymptotically equivalent variation of the moving block bootstrap (MBB) of Künsch (1989) and Liu & Singh (1992). The CBB as a model-free procedure represents a general non-parametric approach that can be used when no distributional assumptions are available and so it shows a behaviour close to the conventional bootstrap resampling. Furthermore, as mentioned earlier, in contrast to MBB, CBB offers the advantage of wrapping data around in a circular before blocking, ensuring that all observations have the same drawing probability. In other words, this method removes the edge effect of MBB by wrapping samples on the circle (for a more detailed survey of block bootstrap methods see Lahiri 1999 and Dudek 2015). The primary challenge in implementing a block bootstrap method lies in selecting an appropriate spatial block size, typically determined empirically (Efron & Tibshirani 1993). The optimal size of blocks varies depending on factors such as the length of the data, the data-generating process, and the specific statistics being considered. To automatically determine the optimal block size for the CBB method, we follow a data-driven approach

proposed by Politis & White (2004) (see also corrections in Patton *et al.* 2009). This method relies on spectral estimation using flat-top lag windows of Politis & Romano (1995). The resampling strategy is summarized by Algorithm 1, while adapted for geophysical applications. Note that an important concept in the sampling process is the randomness in generating each data realization. Consequently, the length of overlap between consecutive blocks is randomly selected during each iteration of random data creation. As a result, at each iteration of the bootstrapping process, a variable number of observed data is provided, depending on the length of overlap. The blocks could be created with fixed or varying lengths at each bootstrap iteration. We purposely opted for a fixed block length, ensuring that the resampled subset of the original data comprises 60 per cent to 70 per cent of all samples. This sampling rate was determined through a sensitivity analysis evaluating the impact of variations in subsample size.

Algorithm 1 CBB resampling algorithm adapted for geophysical applications.

Require: Data $\mathbf{d} \in \mathcal{R}^{m\times 1} = (d_1,\ d_2,\ \ldots,\ d_m)$, and k number of bootstrap iterations, ℓ number of blocks (rows) participating in the construction of bootstrapped samples.

Ensure: Ensemble of bootstrapped data realizations Σ

Initialize: Calculate the optimal block size ζ from Politis & White (2004)
for $t \leftarrow 1$ *to* k **do**
  Draw a uniformly distributed pseudorandom integer $q \sim U(1, \zeta - 1)$, where q represents the length of overlap between consecutive blocks.
  Construct ϑ blocks with $\vartheta \leftarrow (m - \zeta + 1)$ from $\mathbf{d}$ in a circular manner as follows (see Politis & Romano 1992)
  Set $s \leftarrow (\zeta - q)$ and $f \leftarrow \{(\vartheta \times s) + \zeta - 1\}$ ▷ s is the length of shift
  $r \leftarrow 0$
  for $i \leftarrow 1$ *to* f **do**
    $v \leftarrow \left\{i - ((m \times r) - \{i \times \mathrm{fix}\,((m \times r)/i)\})\right\}$ ▷ $\mathrm{fix}(x)$ rounds x to the nearest integer toward zero
    $\mathbf{b}_i \leftarrow d_v$
    if $(i/m) \leftarrow r + 1$ **then**
      $r \leftarrow r + 1$
    end if
  end for
  for $j \leftarrow 1$ *to* ϑ **do**
    $\mathcal{J}_1 \leftarrow \{((j-1) \times s) + 1\}$
    $\mathcal{J}_2 \leftarrow \{((j-1) \times s) + \zeta\}$
    $\mathcal{J} \leftarrow \mathcal{J}_1$ *to* $\mathcal{J}_2$
    Put $\mathbf{B}^*_{j,:} \leftarrow \mathbf{b}_{\mathcal{J}}$
  end for
  where the *jth* block $\mathbf{B}^*_{j\times\zeta}$ with starting point $b_{(j-1)\times s+1}$ consists of ζ elements, i.e., $\mathbf{B}^*_{j,:} \leftarrow \{b_{(j-1)\times s+1},\ b_{(j-1)\times s+2},\ \ldots,\ b_{(j-1)\times s+\zeta}\}$ ▷ $1 \le j \le \vartheta$.
  the matrix $\mathbf{B}^* \in \mathcal{R}^{(\vartheta\times\zeta)}$including the ϑ overlapping subseries (blocks) of length ζ is given as:$\mathbf{B}^* \leftarrow \begin{pmatrix} b_1 & \cdots & b_\zeta \\ \vdots & \ddots & \vdots \\ b_{(\vartheta-1)\times s+1} & \cdots & b_{(\vartheta-1)\times s+\zeta} \end{pmatrix}_{\vartheta\times\zeta}$
  Select randomly ℓ blocks (rows) without replacement from $\mathbf{B}^*$.
  Sort the ℓ blocks (rows) in the ascending order in terms of the number of each row, and then concatenate them together, ensuring that repeated samples are removed, to form the bootstrapped data $\hat{\mathbf{d}} \in \mathcal{R}^{p\times 1}$, where p may vary across each bootstrap iteration.
  Put $\hat{\mathbf{d}}$ into a matrix $\Sigma \in \mathcal{R}^{p\times k}$
end for

2.2 Geostatistical prior modelling

To generate a prior (initial) model using the bootstrapped data $\hat{\mathbf{d}}$, we employ a geostatistical approach based on a 2D ordinary kriging method to approximate the spatial distribution of the prior model cells $\hat{m}_0$. Essentially, by utilizing the spatial correlation information of the resampled data, we can predict the probable variable values at unsampled locations. In geostatistics, the initial stage involves constructing a semivariogram to characterize the spatial correlation of the parameter of interest. Assuming that $\rho_a(x_i, z_i)\ \{i = 1,\ 2,\ \ldots,\ m\}$ represents a generic subsurface sparse point apparent resistivity, the experimental, isotropic semivariogram at a distance $\vec{\lambda} = (\lambda_x, \lambda_z)$ (where $\lambda = \|\vec{\lambda}\|_{l_2}$ is the Euclidean distance between sparse point estimates) in a plane is defined as half of the average squared difference between two data points. This can be expressed as (Stein 1999; Sarma 2010):

$$\psi(\lambda) = \frac{1}{2N(\lambda)} \sum_{i=1}^{N(\lambda)} \left(\rho_a(x_i, z_i) - \rho_a(x_i + \lambda_x, z_i + \lambda_y)\right)^2, \qquad (1)$$

where $N(\lambda)$ is the number of pairs of sparse point estimates between data $\rho_a(x_i, z_i)$ and $\rho_a(x_i + \lambda_x, z_i + \lambda_z)$. The experimental semivariogram serves as an initial estimate for the semivariogram model required for spatial interpolation via kriging. To compute the semivariogram with irregularly spaced data points (as in this study), a search technique is applied following Matheron (1971). By computing the experimental semivariogram, a common semivariogram model such as exponential, spherical, or Gaussian is then fitted to the observed data, represented by variance versus distance. In our study, we found that the stable variogram model, a variant of the Gaussian model, provides the best least-squares fit. This is defined as (Wackernagel 2003):

$$\tilde{\psi}(\lambda) = \mu\left(1 - \exp\left(-\left(\frac{\lambda}{a}\right)^{\alpha}\right)\right) \quad \text{with } 0 < \alpha \le 2 \text{ and } \mu, a > 0, \qquad (2)$$

where μ and a represent the sill and range, respectively. We used the model for a power α equal to 1.5. After obtaining the semivariogram function parameters, ordinary kriging interpolation is applied to the sparse point resistivity values to estimate the subsurface prior resistivity distribution at each model cell $M(p, q)$, where $(q = 1,\ 2,\ \ldots,\ n_x)$ and $(p = 1,\ 2,\ \ldots,\ n_z)$. Here, n_x and n_z are the number of cells in the horizontal and vertical directions, respectively. The advantage of kriging is that it requires no special assumptions regarding the distribution of the parameters (in this case, the measured resistivity values). However, when the spatial distribution of the parameter follows a Gaussian pattern, kriging serves as the best linear unbiased estimator. Thus, it is advisable to apply kriging to variables whose histograms closely resemble a normal distribution (Chilès & Delfiner 2011). Since resistivity data

Table 1. Resistivity distribution of different regions associated with synthetic example 1.

Region	$\rho\ (\Omega \cdot \mathrm{m})$
Topsoil	300
Lens-like body	50
Bedrock	1000

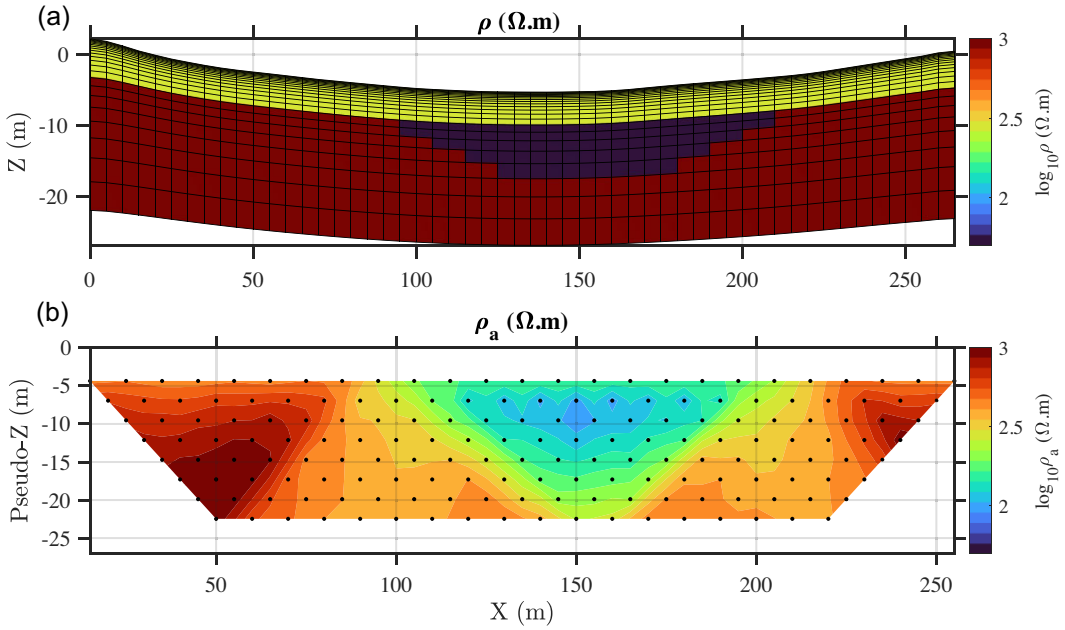

Figure 1. Synthetic Example 1 inspired by a real geological model: (a) true resistivity model featuring a topsoil layer, a conductive lens-like body and a resistive bedrock in the presence of topography. (b) Pseudo-section of apparent resistivity acquired using a collinear Dipole–Dipole array with the horizontal and vertical positions of measurements indicated by black dots.

often follow a lognormal distribution, they are typically transformed into their logarithms (Pryet *et al.* 2011).

2.3 Forward modelling methodology

We start with the assumption that the distribution of the electrical potential, denoted by $u(x, y, z)$, is generated by a single point current source in an isotropic medium characterized by variable conductivity $\sigma(x, y, z)$. Following Ohm's law and the principle of conservation of electric charge, the electrical potential of the point source adheres to the following differential equation, referred to as Poisson's equation, with generalized boundary conditions (Dey & Morrison 1979):

$$-\nabla\cdot(\sigma(x,y,z)\nabla u(x,y,z)) = \nabla\cdot J_s \; in \;\; \Lambda\subset\Re^3. \tag{3}$$

In the above expression, $\nabla\cdot J_s = I\delta(x-x_s)\delta(y-y_s)\delta(z-z_s)$, where J_s is the current density, I indicates the input electrical current, and δ is Dirac's delta function, Λ is the computational domain and x_s, y_s and z_s are the point source coordinates.

Assuming that the y direction is parallel to the strike direction of the model, eq. (3) reduces, in Cartesian coordinates, to

$$-\nabla\cdot(\sigma(x,z)\nabla u(x,y,z)) = I\delta(x-x_s)\delta(y-y_s)\delta(z-z_s). \tag{4}$$

In practice, to take advantage of the 3D current source characteristics, a spatial Fourier-Cosine transform of the potential $u(x, y, z, \omega)$ for a sufficiently large number of discrete wavenumbers (k_y) is carried out along the medium invariant y-direction. By applying the Fourier-Cosine transform to eq. (4), we arrive at the following 2D Helmholtz equation:

$$\begin{aligned}\partial_x\ &\left(\sigma(x,z)\,\partial_x\widetilde{u}\left(x,k_y,z\right)\right) - k_y^2\sigma(x,z)\widetilde{u}\left(x,k_y,z\right) \\ +\partial_z\ &\left(\sigma(x,z)\,\partial_z\widetilde{u}\left(x,k_y,z\right)\right) = -\tfrac{1}{2}I\delta(x-x_s)\,\delta(z-z_s),\end{aligned} \tag{5}$$

where $\widetilde{u}\left(x,k_y,z\right)$ is the transformed electrical potential in the Fourier domain. Eq. (5) serves as the governing equation for 2.5D direct-current resistivity modelling. By applying a homogeneous Neumann boundary condition at the ground interface and a homogeneous mixed boundary condition at non-air boundaries to solve eq. (5), a 2D distribution of electrical potential can be obtained for a range of wavenumbers. To obtain potential solutions in the spatial domain, an inverse Fourier-Cosine transform is then utilized. The integration of eq. (5) is numerically carried out using a finite-difference algorithm, following the discretization-by-area approach established by Dey & Morrison (1979). The computational domain is divided into quadrilateral meshes, with finer discretization applied to the foreground and coarser discretization to the background. The foreground section consists of two cells (nodes) positioned between adjacent electrodes horizontally, while the depth of these cells increases logarithmically in the vertical direction. Conversely, cell sizes in the background region expand gradually away from the foreground to accommodate far-field boundary conditions. Furthermore, To address surface topography, we employ an efficient mapping technique called the Schwarz–Christoffel transformation approach. This method converts the problem from a plane with non-uniform topography (W-plane) into a plane with flat topography (Z-plane). By utilizing the Schwarz–Christoffel transformation parameters obtained from the solution process (as described in Chuang *et al.* 1993), along with the inverse Schwarz–Christoffel transformation (Costamagna 1987), we can adjust the model cells vertically to accommodate the topographic irregularities. The principles and applications of the Schwarz–Christoffel transformation in electrical resistivity tomography are thoroughly discussed in Ghanati & Fallahsafari (2023).

2.4 Inverse modelling methodology

The inverse problem that we address involves estimating the resistivity (or equivalently, conductivity) distribution of the Earth's unknown subsurface model given measurements taken at the surface. This problem can be formulated as a nonlinear inverse problem

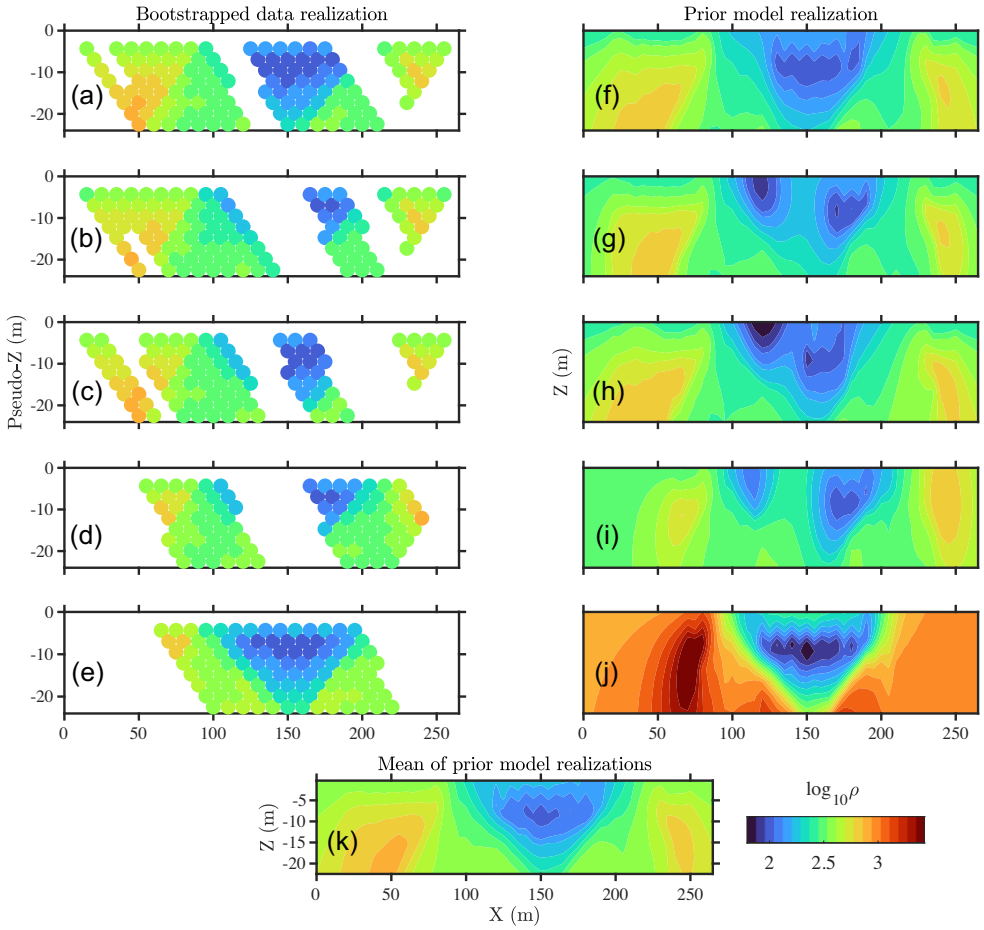

Figure 2. Example realizations of bootstrapped samples (a)–(e) with corresponding prior models generated using the geostatistical method (f)–(j), and the mean of prior model realizations (k) for synthetic example 1.

Table 2. Total CPU time (in hour) required for implementing the proposed uncertainty quantification method across synthetic and real examples. It is important to note that following discretization, the vertical and horizontal dimensions of the model grid ($N_z \times N_x$) for synthetic examples 1 and 2, as well as real cases 1 and 2, are 30×65, 35×95, 30×40 and 30×40, respectively.

Synthetic example 1	Synthetic example 2	Real case 1	Real case 2
4.2	16.5	35.8	18.7

which theoretically relates the observations $d \in \mathcal{R}^{m\times 1}$ to the subsurface true conductivity $(\sigma)\, m \in \mathcal{R}^{n\times 1}$ using a parametric statistical model of form,

$$d = \mathcal{F}(m) + \mathcal{V}, \tag{6}$$

where $\mathcal{F}: \mathcal{R}^{n\times 1} \rightarrow \mathcal{R}^{m\times 1}$ is the mapping function and $\mathcal{V} \in \mathcal{R}^{m\times 1}$ is the measurement data error. Assuming that the data error $\mathcal{V}$ is a Gaussian distributed random vector with mean zero and covariance matrix $\mathcal{D}_\delta \in \mathcal{R}^{m\times m}$, i.e., $\mathcal{V} \sim \mathcal{N}(0, \mathcal{D}_\delta)$, the likelihood function is defined as:

$$p(d|m) \propto \frac{1}{\sqrt{(2\pi)^m |\mathcal{D}_\delta|}} \exp\left(-\frac{1}{2}\left\|\mathcal{D}_\delta^{-1/2}(d - \mathcal{F}(m))\right\|_{l_2}^2\right), \tag{7}$$

where $|\mathcal{D}_\delta|$ represents the determinant of the covariance matrix $\mathcal{D}_\delta$. Maximizing the likelihood function is equivalent to the minimization of a misfit function that has the following canonical form:

$$\mathcal{Q}(m) = arg \min_{m\in\mathcal{R}^{n\times 1}} \frac{1}{2}\left\|\mathcal{D}_\delta^{-1/2}(d - \mathcal{F}(m))\right\|_{l_2}^2. \tag{8}$$

The solution obtained by minimizing eq. (8) tends to be unstable and not physically realistic, primarily because of errors in the observed data and the ill-posed nature of the forward operator matrix. To address this issue, one approach is to introduce a suitable stabilizing function into the minimization problem. This additional term is designed to incorporate prior knowledge into the inversion process, aiding in the filtration of undesired features and constraining the model within desired spaces. Therefore, we propose a regularized Homotopy continuation method, which either mitigates the numerical impact of ill-conditioning or enhances the global convergence properties of the inversion algorithm. The fundamental concept behind the homotopic method, a traditional element of topology, is to transform the original problem into a simpler one with a known solution. Subsequently, the process involves continuously deforming the simpler problem back into the original one, while monitoring the sequence of zeros that connect the solution of the simpler problem to that of the original problem. The fundamental concept of homotopic optimization is broadly outlined in Ghanati & Müller-Petke (2021) as well as Roudsari *et al.* (2024). For the sake of

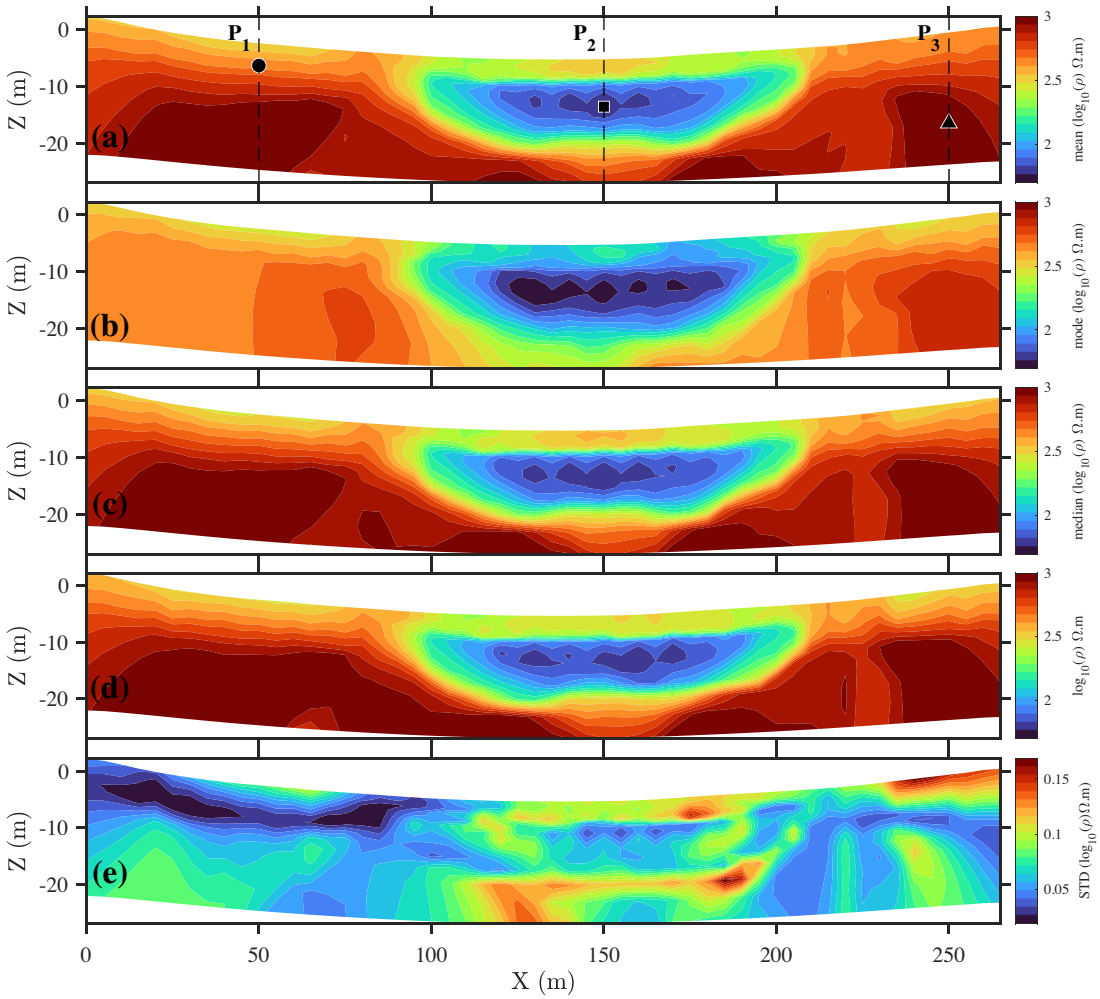

Figure 3. Statistical metrics derived from the inversion of the synthetic model in Fig. 1 using the proposed algorithm: (a) arithmetic mean, (b) mode (most frequent value), (c) median, (d) inverted section from the unbootstrapped (original) data, and (e) standard deviation. Vertical dashed lines mark the locations of the posterior resistivity profiles displayed in Fig. 5.

this discussion, we will succinctly recapitulate the key principles here.

Let $\mathcal{J}: \mathcal{X} \rightarrow \mathcal{Y}$ be a nonlinear convex operator between Hilbert spaces $\mathcal{X} \in \mathcal{R}^n$ and $\mathcal{Y} \in \mathcal{R}^n$. We are interested in the problem of the form

$$\mathcal{J}(X) = 0, \tag{9}$$

where X represents the unknown parameters $X = \{x_1, x_2, \ldots, x_n\}$ in a set of nonlinear equations. Without an appropriate approximation of the zero point, attempting an iterative solution such as the conventional Newton method for eq. (9) is prone to failure, as it is probable that inadequate initial values will be selected. As a remedy, we can turn the problem into a homotopic continuation function $\mathcal{H}(X, \mathcal{T})$:

$$\mathcal{H}(X, \mathcal{T}) \equiv \mathcal{C}(\mathcal{T})\mathcal{J}(X) + \mathcal{G}(\mathcal{T})\mathcal{L}(X) = 0. \tag{10}$$

In this equation, $\mathcal{H}: \mathcal{R}^n \times \mathcal{R} \rightarrow \mathcal{R}^n$, $\mathcal{J}(X)$ indicates target function, $\mathcal{L}(X)$ is auxiliary homotopy function that must be known or controllable and easy to solve, and $\mathcal{T} \in [0, 1]$ is embedding or continuation parameter such that $\mathcal{C}(0) = \mathcal{G}(1) = 0$ and $\mathcal{C}(1) = \mathcal{G}(0) = 1$ resulting in, respectively, $\mathcal{H}(X, 0) = \mathcal{L}(X)$ and $\mathcal{H}(X, 1) = \mathcal{J}(X)$, and $\mathcal{C}(\mathcal{T}) > 0$, $\mathcal{G}(\mathcal{T}) > 0$ whenever $0 < \mathcal{T} < 1$. By substituting $\mathcal{C}(\mathcal{T})$ and $\mathcal{G}(\mathcal{T})$ with $\mathcal{T}$ and $(1 - \mathcal{T})$, we rewrite eq. (10) as

$$\mathcal{H}(X, \mathcal{T}) \equiv \mathcal{T}\mathcal{J}(X) + (1 - \mathcal{T})\mathcal{L}(X) = 0. \tag{11}$$

Our goal is to solve the homotopy function $\mathcal{H}(X, \mathcal{T}) = 0$ instead of $\mathcal{J}(X) = 0$ by monotonically increasing the embedding parameter $\mathcal{T}$ from 0 to 1 as the auxiliary function $\mathcal{L}(X) = 0$ is continuously deformed to the target problem $\mathcal{J}(X) = 0$. In the context of a least-squares problem, by substituting $\mathcal{J}(X)$ and $\mathcal{L}(X)$ with d and $\mathcal{F}(m_0)$, respectively, eq. (11) is formulated as

$$\mathcal{H}(m, \mathcal{T}) \equiv \mathcal{T}d + (1 - \mathcal{T})\mathcal{F}(m_0), \tag{12}$$

where d indicates the measured data and $\mathcal{F}(m_0)$ is the forward response of the problem in terms of the initial solution. To trace the Homotopy curve numerically $\mathcal{T} = 0$ to $\mathcal{T} = 1$, we utilize a continuation scheme that evenly divides the interval $[\mathcal{T} = 0, \mathcal{T} = 1]$ into $\mathcal{G}$ embedding parameters (see Ghanati & Müller-Petke 2021). Remembering the importance of the stabilizing functional and substituting eq. (12) into eq. (8), the objective function in eq. (8) is

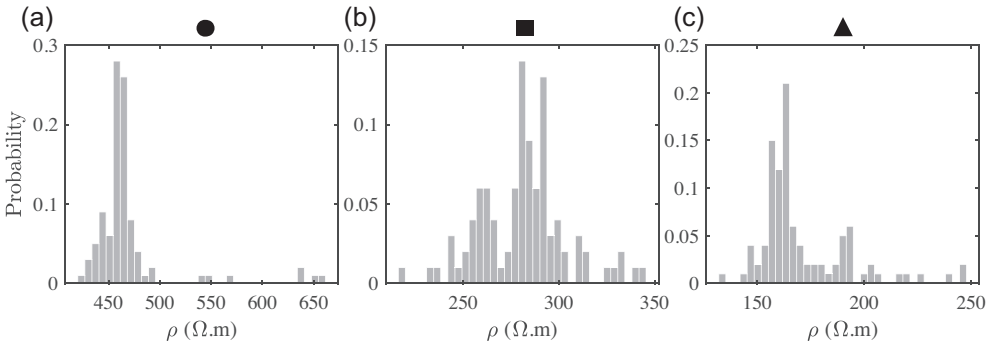

Figure 4. Posterior distributions of resistivity for the points depicted in Fig. 3, represented by (a) circle, (b) square, and (c) triangle.

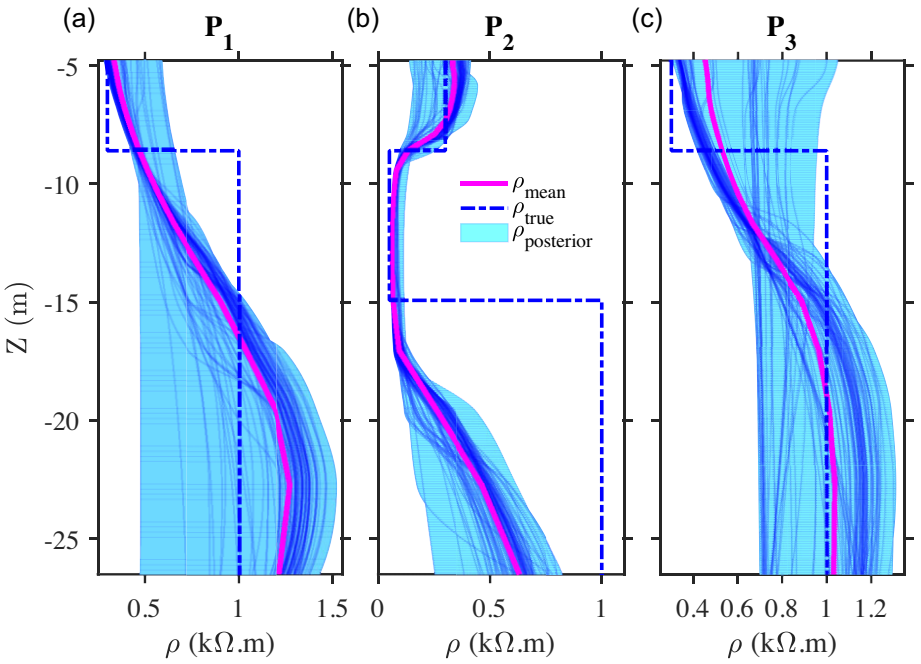

Figure 5. Vertical resistivity profiles along three transects (P_1, P_2 and P_3) shown in Fig. 3, obtained using inversion of the bootstrapped data. The shaded area indicates the inferred posterior resistivity distribution, the solid line represents the mean model and the dashed line depicts the target model as a function of depth.

structured as an unconstraint optimization problem:

$$\widetilde{\mathcal{Q}}^{\mathfrak{J}}(m) = arg \min_{m\in\mathcal{R}^{n\times 1}} \left\{ \frac{1}{2} \left\| \mathcal{D}_{\delta}^{-1/2} \left(\mathcal{H}(m,\mathcal{T}) - \mathcal{F}(m) \right) \right\|_{l_2}^{2} + \frac{\mathfrak{J}}{2} \left\| \Gamma \left(m - m_{apr} \right) \right\|_{q}^{q} \right\}. \tag{13}$$

Or equivalently

$$\widetilde{\mathcal{Q}}^{\mathfrak{J}}(m) = arg \min_{\mathbf{m}\in\mathcal{R}^{n\times 1}} \left\{ \frac{1}{2} \left\| \mathcal{D}_{\delta}^{-1/2} \left(\mathcal{T}d + (1-\mathcal{T})\mathcal{F}(m_0) - \mathcal{F}(m) \right) \right\|_{l_2}^{2} + \frac{\mathfrak{J}}{2} \left\| \Gamma \left(m - m_{apr} \right) \right\|_{q}^{q} \right\}. \tag{14}$$

In the given equation, the term $\left\| \Gamma \left(m - m_{apr} \right) \right\|_{q}^{q}$ represents the stabilizing functional which is similar to the prior function $p(m) \propto \exp\left(-\frac{\mathfrak{J}}{2} \left\| \Gamma \left(m - m_{apr} \right) \right\|_{q}^{q} \right)$ in the Bayesian inversion setting. The stabilizing functional takes to account the distance to the prior model m_{apr}, where $q > 0$ and $\|\mathbf{x}\|_q^q = \sum_i |x_i|^q$ stands for the l_q-(quasi) norm of the vector $\mathbf{x}$, $\mathfrak{J} \in \mathcal{R}^+$, a set of real positive values, is the Lagrangian multiplier, which controls the balance between the data-fidelity and stabilizing term, and the matrix $\Gamma = (a_x \Gamma_x^T \Gamma_x + a_z \Gamma_z^T \Gamma_z)$ serves as the constraint matrix, where it is defined by the discrete partial first or second-order derivatives in both the horizontal and vertical directions. The coefficients $a_x, a_z \in \mathcal{R}^+$ are employed to enhance smoothness in either of these directions. The minimization of $\widetilde{\mathcal{Q}}^{\mathfrak{J}}(m)$ is significantly influenced by the parameter q. Setting $q = 2$ leads to the generation of smooth and small models, whereas setting $q = 1$ results in sparse and blocky models.

Assuming $q = 2$ and linearizing $\mathcal{F}(m)$ via the computation of Fréchet derivatives ($J = \partial\mathcal{F}(m)/\partial m$; see Ghanati & Fallahsafari 2022 for details), the solution of eq. (14) leads to the following iterative numerical method, i.e.,

$$\begin{aligned} \delta m_p = \mu_p \left\{ J^T \left(m_{p-1} \right) \mathcal{D}_{\delta}^T \mathcal{D}_{\delta} J \left(m_{p-1} \right) + \mathfrak{J}_p (a_x \Gamma_x^T \Gamma_x + a_z \Gamma_z^T \Gamma_z) \right\}^{-1} \\ \times \left\{ J^T \left(m_{p-1} \right) \mathcal{D}_{\delta}^T \mathcal{D}_{\delta} \left[\mathcal{T}_p d + (1-\mathcal{T}_p) \mathcal{F}(m_0) - \mathcal{F}(m_{p-1}) \right] \right. \\ \left. - \mathfrak{J}_p (a_x \Gamma_x^T \Gamma_x + a_z \Gamma_z^T \Gamma_z)(m_{p-1} - m_{apr}) \right\}, \end{aligned} \tag{15}$$

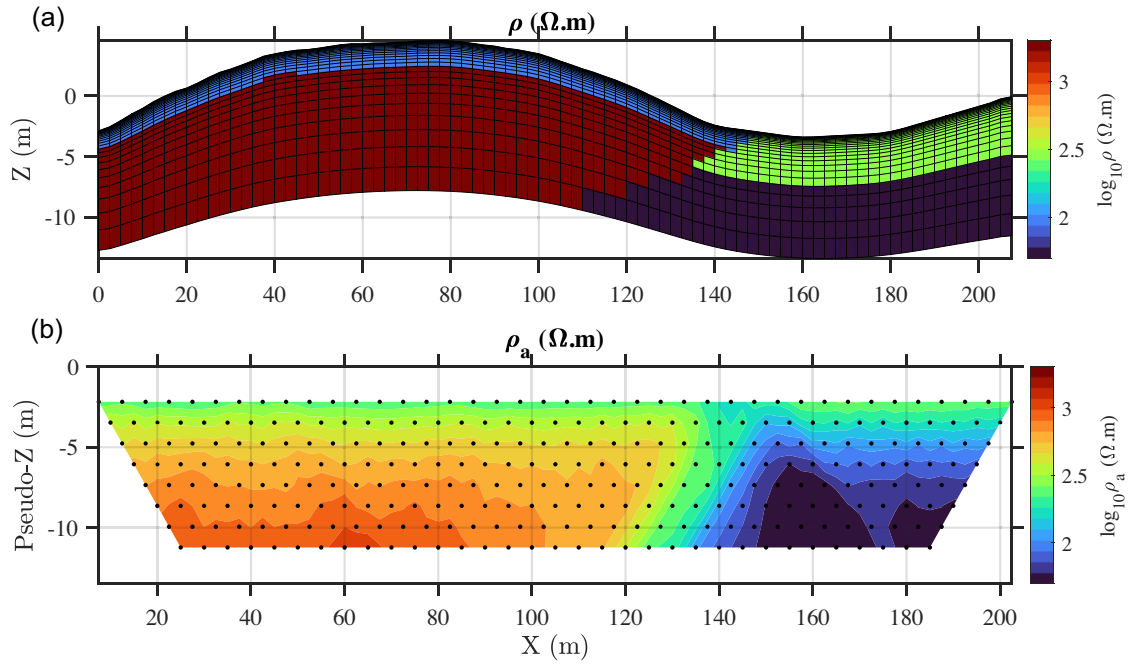

Figure 6. Synthetic example 2 inspired by a real geological model: (a) true resistivity model depicting a topsoil layer and a resistive bedrock on the left side of the fault, and a surface resistive layer and a conductive bedrock on the right side of the fault, considering topography. (b) Pseudo-section of apparent resistivity acquired using a collinear Dipole–Dipole array, with the horizontal and vertical positions of measurements denoted by black dots.

where the parameter μ_p denotes the step length, taking on a positive value within the range $(0,\ 1]$ to ensure a reduction in the global objective function across successive iterations.

When using the natural logarithm of the model parameters (resistivity or equivalently conductivity), the residual vector $\delta d = \mathcal{T}_p d + (1-\mathcal{T}_p)\mathcal{F}(\mathbf{m_0}) - \mathcal{F}(m_{p-1})$ needs to consider this adjustment:

$$\delta d = \log_e\left(\mathcal{T}_p d + (1-\mathcal{T}_p)\mathcal{F}(\mathbf{m_0})\right) - \log_e\left(\mathcal{F}(m_{p-1})\right). \tag{16}$$

Eq. (15) can be solved by applying the Gauss–Newton iterative scheme, which leads to smooth solutions.

By setting $q = 1$, the system of equations to be solved is therefore:

$$\begin{aligned}\delta m_p^{\iota} = \mu_p \{ & J^T(m_{p-1})\mathcal{D}_\delta^T\mathcal{D}_\delta J(m_{p-1}) + \mathfrak{J}_p(a_x\Gamma_x^T\Psi(\delta m_p^{\iota})\Gamma_x + a_z\Gamma_z^T\Psi(m_{p-1})\Gamma_z)\}^{-1} \\ & \times \{J^T(m_{p-1})\mathcal{D}_\delta^T\mathcal{D}_\delta\left[\mathcal{T}_p d + (1-\mathcal{T}_p)\mathcal{F}(\mathbf{m_0}) - \mathcal{F}(m_{p-1})\right] \\ & -\mathfrak{J}_p(a_x\Gamma_x^T\Psi(\delta m_p^{\iota})\Gamma_x + a_z\Gamma_z^T\Psi(m_{p-1})\Gamma_z)(m_{p-1} - m_{apr})\},\end{aligned} \tag{17}$$

where

$$\Psi(\delta m_p^{\iota}) = diag_{i=1,\ldots,n}\left(\left(\Gamma_x\delta m_p^{\iota}\right)^2 + \left(\Gamma_z\delta m_p^{\iota}\right)^2 + \varsigma\right)^{-1/2}. \tag{18}$$

Here, $\Psi(\delta m_p^{\iota})$ represents the diagonal weighting matrix, where p and ι denote the iteration indices for the regularized homotopy step and the iteratively reweighted least-squares (IRLS) iteration, respectively. $\varsigma \in \mathcal{R}^+$ represents a small positive value. We employ the IRLS method to solve eq. (17). The IRLS algorithm, aimed at finding a minimizer of $\widetilde{\mathcal{Q}}^{\mathfrak{J}}(m)$ when q is set to 1, is outlined in Algorithm 2. Furthermore, the penalty parameter $\mathfrak{J}_p$ is updated at each iteration of the homotopic continuation inversion according to the guideline introduced by Roudsari *et al.* (2024). In Algorithm 2, upper and lower bounds are imposed to prevent unrealistic solutions by using the constraint conditions defined in line 13.

Algorithm 2 IRLS-based homotopic continuation inversion algorithm (for the definitions of the symbols, see the text).

Require: $\mathbf{d}, \mathbf{m}_0, \mathcal{D}_\delta, \mathcal{T} \in [0,\ 1]$ ($\mathcal{T}$ is evenly divided into $\mathcal{G}$ spaces), ϵ, Γ_x, Γ_z, η (lower boundary of $\mathbf{m}$), ϑ (upper boundary of $\mathbf{m}$), a_x, and a_z

1: **Initialize:** Set $\varsigma \leftarrow 1e-6$, $\Psi \leftarrow I_n$ (Identity matrix), and $\mathfrak{J}_0 \leftarrow 1e-6$
2: **for** $p \leftarrow 1$ *to* $\mathcal{G}$ **do**
3: **Calculate** homotopy function $\mathcal{H}\left(\mathbf{m}, \mathcal{T}_p\right) = \mathcal{T}_p\mathbf{d} + \left(1-\mathcal{T}_p\right)\mathcal{F}\left(\mathbf{m}_0\right)$
4: $\iota \leftarrow 1$
5: Set the model update $\delta\mathbf{m}^{\iota-1} \leftarrow 0$ and $\delta\mathbf{m}^{\iota} \leftarrow 1$ to enter the inner loop
6: **while** $\|\delta\mathbf{m}^{\iota} - \delta\mathbf{m}^{\iota-1}\|_{l_2} > \epsilon$ or *maximum number of iterations is not met* **do**
7: Calculate $\delta\mathbf{m}^{\iota}$ using eq 17
8: Update Ψ using eq 18
9: $\iota \leftarrow \iota + 1$
10: **end while**
11: Implement line search to find μ_p
12: Update $\delta\mathbf{m}_p \leftarrow \mu_p\delta\mathbf{m}^{\iota}$
13: Impose constraint conditions to obtain $\mathbf{m}_p$ using$\mathbf{m}_p \leftarrow \frac{[\vartheta(\mathbf{m}_{p-1}-\eta)\exp(\delta\mathbf{m}_p)+\eta(\vartheta-\mathbf{m}_{p-1})]}{[(\mathbf{m}_{p-1}-\eta)\exp(\delta\mathbf{m}_p)+(\vartheta-\mathbf{m}_{p-1})]}$
14: Update $\mathfrak{J}_p$ using the algorithm proposed by Roudsari *et al.* (2024)
15: **end for**

Ensure: $\sigma \leftarrow \mathbf{m}_p$

2.5 Algorithm implementation

After establishing the framework for the inversion process, the subsequent stage involves detailing the implementation of the proposed uncertainty quantification method. The overall algorithm comprises three main steps: (1) perturbing the observed data using the CBB

Table 3. Resistivity distribution of different regions associated with synthetic example 2.

Region	ρ ($\Omega \cdot$ m)
Overburden layer (topsoil)	100
Resistive bedrock	1000
Surface resistive layer	500
Conductive bedrock	50

algorithm to generate a randomly bootstrapped data set $\hat{\mathbf{d}}$, where the observations $\hat{\mathbf{d}}$ are treated as a realization of random vector, (2) stochastically generating an initial model $\hat{\mathbf{m}}_0$ for the corresponding bootstrapped data set using the geostatistical method and (3) obtaining the inverted model using the homotopic continuation method. It is noteworthy that perturbing the data and generating the initial model transforms the deterministic optimization problem in eq. (14) into a stochastic optimization problem, allowing for exploring and sampling the different models of the equivalence region. With the above conditions, eq. (14) turns into the following form

$$\widetilde{\mathcal{Q}}^{\mathfrak{I}}(\hat{\mathbf{m}}) = arg \min_{\hat{\mathbf{m}} \in \mathcal{R}^{n\times 1}} \left\{ \frac{1}{2} \left\| \mathcal{D}_{\delta}^{-1/2} \left(\mathcal{T}\hat{\mathbf{d}} + (1-\mathcal{T})\mathcal{F}(\hat{\mathbf{m}}_0) - \mathcal{F}(\hat{\mathbf{m}}) \right) \right\|_{l_2}^2 + \frac{\mathfrak{I}}{2} \left\| \Gamma \left(\hat{\mathbf{m}} - \hat{\mathbf{m}}_{apr} \right) \right\|_q^q \right\}. \quad (19)$$

As the optimization process in eq. (19) is stochastic in nature, its solution ($\hat{\mathbf{m}}$) is also stochastic. Moreover, since there is generally no prior information about the prior or reference model in real case scenarios, we assign $\hat{\mathbf{m}}_{apr}$ to $\hat{\mathbf{m}}_0$. Furthermore, while the inverted models of all bootstrapped data sets are computed independently, it is feasible to parallelize the for-loop when sufficient resources are accessible.

The proposed algorithm's implementation is found to be quite straightforward and easily integrates with any regularized inversion method, requiring minimal parameter tuning before applying the resampling algorithm. One of these parameters is the choice of the block length that needs to be defined by the user. At each bootstrap iteration, the blocks can have fixed or adjustable lengths. To maintain consistency, we deliberately chose a fixed block length, resulting in the resampled subset of the original data containing 60 per cent to 70 per cent of all samples. As alluded to earlier, employing partial information results in ambiguity within the inversion outcomes, enabling the sampling of equivalent solutions. In the context of mathematics, uncertainty analysis in discrete inverse problems entails discovering an ensemble of models that are consistent with the prior information and adequately fit the data within the same error bound δ (Grayver *et al.* 2016; Fernández-Muñiz *et al.* 2019):

$$M_e = \left\{ \hat{\mathbf{m}} : \ \| \hat{\mathbf{d}} - \mathcal{F}(\hat{\mathbf{m}}) \|_{l_2} < \delta \right\}, \quad (20)$$

where the uncertainty space is defined by considering the equivalent models (M_e) that meet the above condition for a given noise level. Algorithm 3 summarizes different steps of the implementation of the proposed uncertainty quantification method. It is essential to note that Algorithm 3 does not require any extra adjustments beyond those necessary to execute a single deterministic inversion.

Algorithm 3 Sketch of the proposed uncertainty quantification repeatedly executed to minimize a stochastic cost function.

Require: observed data $\mathbf{d}$, number of iterations k
Initialize k processes
Ensure:
1: Draw a bootstrapped data realization $\hat{\mathbf{d}}$ using Algorithm 1
2: Draw an initial (prior) model $\hat{\mathbf{m}}_0$ corresponding to the bootstrapped data using the geostatistical method
3: Minimize eq.19 to find the model parameters $\hat{\mathbf{m}}$
4: Check weather $\hat{\mathbf{m}}$ meet the condition given by eq.20, otherwise go to line 1
5: Put $\hat{\mathbf{m}}$ into the matrix $\mathbf{M}_e$ to construct the posterior distribution of the model parameters

3 NUMERICAL EXPERIMENTS

In this section, we present a set of experiments using two synthetic examples and two real data sets with varying geological conditions, all of which include topographical variations. Our focus is on assessing the effectiveness of the proposed algorithm in quantifying the variability in model predictions during the regularized inversion of ERT data, considering the influence of uncertainty in observed data and initial model assumptions. It is noteworthy that the numerical examples provided here primarily consist of nonlinear cases, due to the inherent complexity of uncertainty analysis in such inverse problems. Nevertheless, the proposed strategy can be conveniently adapted and extended to accommodate linear problems as well. It is also noted that in both synthetic and real data inversion, the weighting coefficients a_x and b_x for horizontal and vertical boundaries were set to one. All numerical computations are conducted on a desktop computer equipped with a 3.6 GHz dual-core Intel processor. Due to the independence of each bootstrapped data realization's inversion, it is possible to distribute computations across different central processing units (CPUs) using the parfor command in MATLAB. This approach can significantly reduce processing time, especially on multicore machines.

3.1 Theoretical examples

We begin with the first synthetic example, which consists of an overburden layer and a resistive background containing a conductive lens-like body intended to mimic a freshwater lens in a sedimentary environment. Table 1 represents different regions of the first example in terms of resistivity values. Fig. 1(a) shows the discretization of the true synthetic model in the presence of topography. It is assumed that 28 electrodes with a minimum electrode spacing of 10 *m* are deployed at the surface. A Dipole–Dipole scheme is conducted to generate the apparent resistivity values up to eight levels leading to a total of 172 measurements. The simulated data are perturbed with uncorrelated Gaussian noise using the standard deviation of 5 per cent of each data amplitude. This represents the typical order of data discrepancies calculated through the inversion of ERT field data. Note that in both synthetic examples, the forward modelling mesh is different from the inversion mesh to avert an inverse crime (Kaipio & Somersalo 2007). Fig. 1(b) shows the pseudo-section of apparent resistivity derived from the forward calculation of the first synthetic model with surface topography. The apparent resistivity pseudo-section indicates the expected presence of the conductive anomaly compared to the background medium in the middle of the pseudo-section.

To obtain the posterior distribution of the model parameters, the proposed strategy described in Algorithm 3 is performed for 100 iterations. In Fig. 2, we display several randomly selected realizations of the bootstrapped data alongside the corresponding prior models generated by the geostatistical method.

All bootstrapped data realizations are inverted in the context of stochastically perturbing the observed data and prior model using

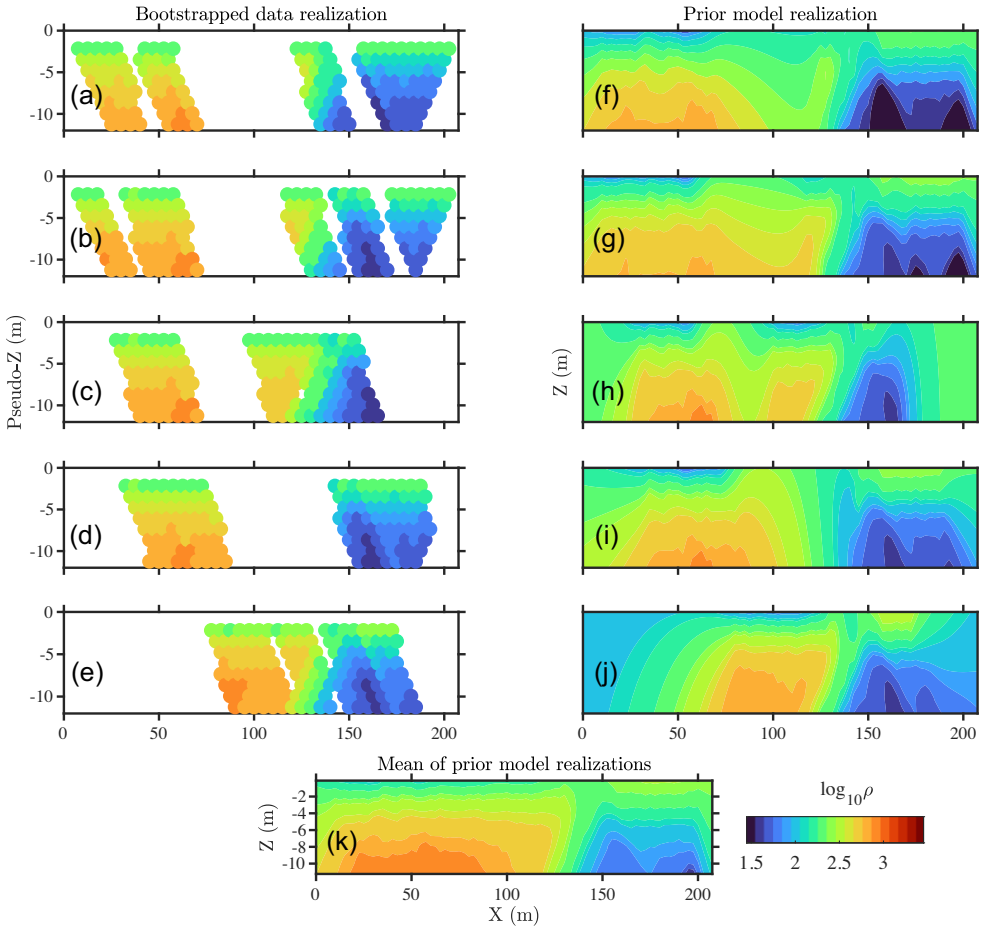

Figure 7. Example realizations of bootstrapped samples (a)–(e) with corresponding prior models generated using the geostatistical method (f)–(j), and the mean of prior model realizations (k) for synthetic example 2.

consistent inversion parameters resulting in multiple random solutions that adequately fit the resampled data. In essence, as per eq. (20), solutions that fulfill the given condition are considered equivalent models. To enhance the resolution of sharp boundaries and blocky features, the inversion process is carried out using Algorithm 2, which involves setting $q = 1$ in eq. (19). It is worth noting that the inversion of each data realization converges to the final solution in 10 iterations according to the number of continuation parameters ($\mathcal{T}$). The CPU time required for implementing the entire inversion is reported in Table 2. Figs 3(a)–(c) display various statistical measures, such as the arithmetic mean, mode (i.e. maximum frequent value) and median, derived from the inversion of all bootstrapped data realizations. These measures provide insight into the estimation of the true resistivity distribution. The analysis reveals that the topsoil layer and the resistive background are relatively well resolved throughout the model. However, the recovery of the lens-like body varies across the different statistics. For comparison, the inverted resistivity tomogram of the unbootstrapped (original) data is shown in Fig. 3(d), where only trivial differences are observed between this section and, in particular, the mean and median sections. Furthermore, Fig. 3(e) illustrates the uncertainty map of the solution, presented in terms of standard deviation. As expected, elevated uncertainty is evident along the edges of the conductive anomaly. Moreover, uncertainty increases in areas characterized by inadequate data coverage. Significant uncertainties also arise between the overburden layer and the conductive anomaly. This arises from the electrical current's propensity to traverse the resistive layer, resulting in a weakened signal from the resistive layer.

To evaluate the recovery of the resistivity model, we choose three points representing the topsoil layer (circle), lens-like body (square) and bedrock (triangle). Fig. 4 displays the posterior distribution of resistivity values corresponding to these locations. Regarding the resolved parameters, the distributions exhibit an approximate bell shape with a solitary peak. Nonetheless, certain marginal distributions display slight skewness either to the right or left.

We also provide a direct comparison of the resistivity distributions, the mean model, and the true model along three vertical profiles marked by black dashed lines in Fig. 5. This comparison shows a relatively good agreement between the mean model (magenta line) and the true model (blue line). Nevertheless, the variations in resistivity profiles tend to escalate with depth, attributed to the diminishing resolving capability of electrical resistivity tomography data as depth increases.

The second synthetic model incorporates a more intricate scenario, drawing inspiration from a genuine geological model. It features an inclined fault dividing the model into two distinct media. The left medium comprises a conductive topsoil overlying a resistive bedrock, while the right medium consists of a resistive surface layer resting on a highly conductive medium. Additionally, the model

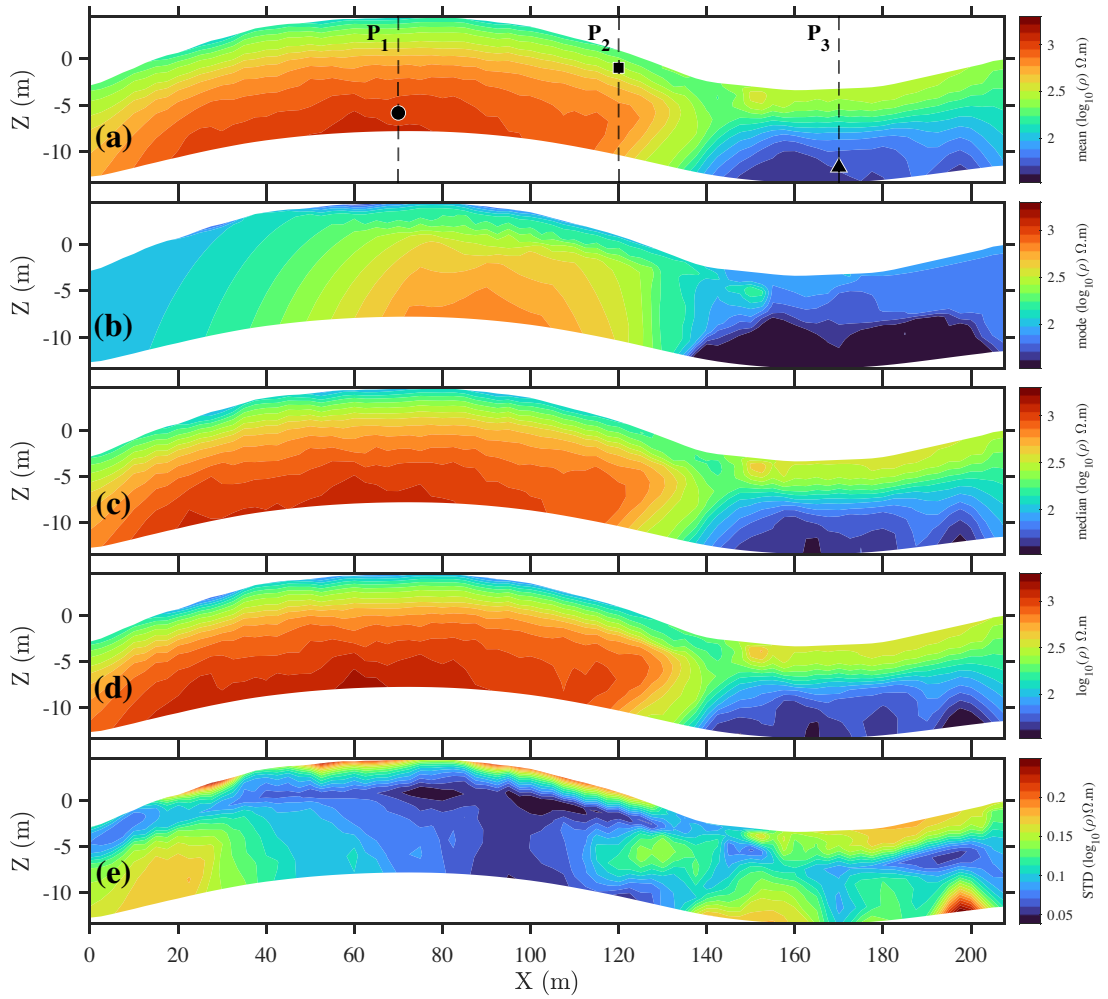

Figure 8. Statistical metrics derived from the inversion of the synthetic model in Fig. 6 using the proposed algorithm: (a) arithmetic mean, (b) mode (most frequent value), (c) median, (d) inverted section from the unbootstrapped (original) data, and (e) standard deviation. Vertical dashed lines mark the locations of the posterior resistivity profiles displayed in Fig. 10.

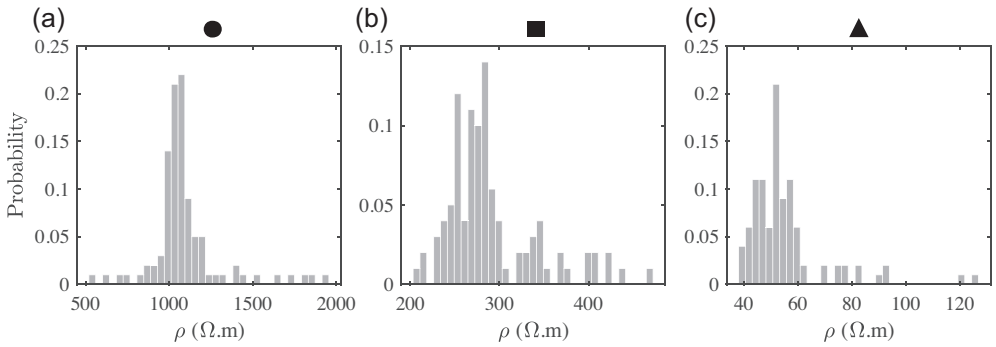

Figure 9. Posterior distributions of resistivity for the points depicted in Fig. 8, represented by (a) circle, (b) square and (c) triangle.

considers irregular terrain (Fig. 6a). To generate synthetic apparent resistivity data, a collinear Dipole–Dipole arrangement is utilized, comprising 43 electrodes evenly spaced at 5 *m* intervals, yielding a total of 292 data points. The current and potential dipoles maintain an interval separation ranging from 5 to 40 *m*. This procedure is carried out on the simulated Earth model, incorporating the resistivity values listed in Table 3. To improve realism, Gaussian noise with an uncorrelated distribution and a standard deviation equivalent to 5 per cent of each data amplitude was incorporated into the data set. In Fig. 6(b), the forward response is presented as a pseudo-section, highlighting that the two distinct mediums resulting from the fault can still be differentiated.

Following the strategy used for the first synthetic example, the proposed strategy described in Algorithm 3 is executed for 100 iterations. We present a few randomly selected example realizations of the resampled data along with the corresponding

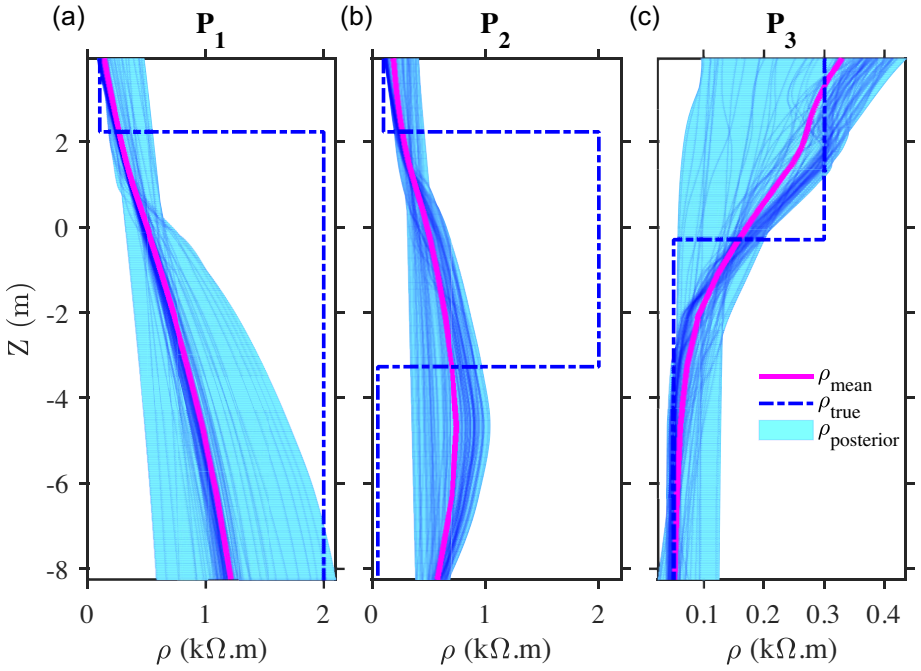

Figure 10. Vertical resistivity profiles along three transects (P_1, P_2 and P_3) shown in Fig. 8, obtained using inversion of the bootstrapped data. The shaded area indicates the inferred posterior resistivity distribution, the solid line represents the mean model and the dashed line depicts the target model as a function of depth.

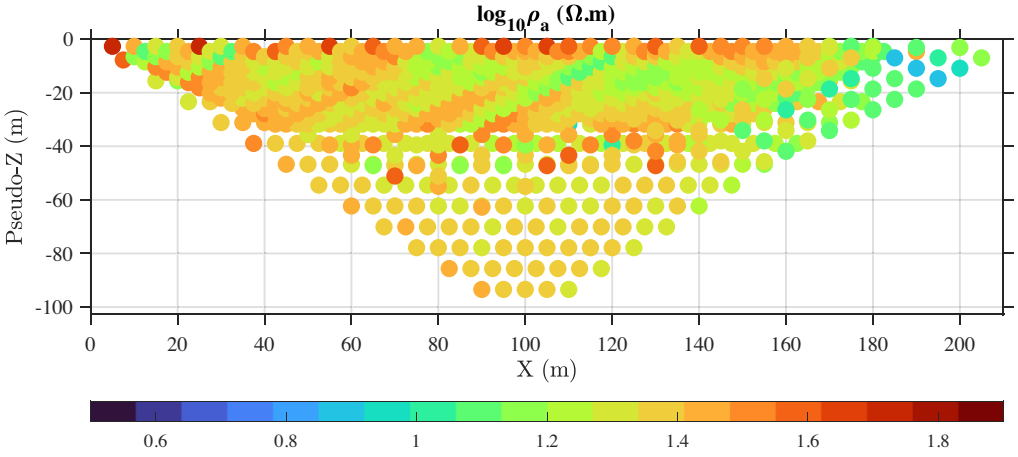

Figure 11. Resistivity pseudo-section associated with real case 1, acquired using a combination of Dipole–Dipole and Wenner-alpha arrays with minimum electrode spacings of 5 and 10 *m*, respectively.

prior models generated by the proposed geostatistical method in Fig. 7.

Similar to the synthetic example 1, the inversion of each bootstrapped data realization reaches convergence after 10 iteration. Table 2 reports the CPU time needed to execute the entire inversion. The resulting solution ensemble is then analysed using various statistical metrics, such as the arithmetic mean, mode and median. Figs 8(a)–(c) illustrate the distribution of resistivity in the subsurface model, as represented by these ensemble metrics. It can be observed that the true resistivity structures, particularly the boundary between the underlying layers and the surface layers, are effectively represented by each of these descriptive statistics, with the exception of the mode section to some extent. The greater discrepancy of the mode section compared to the mean and median sections is due to the fact that the mode is defined as the most frequent value and is not influenced by all the posterior model realizations. However, the fault geometry does not appear to be perfectly resolved in all statistical metrics. For comparison, Fig. 8(d) presents the inverted resistivity tomogram derived from the unbootstrapped (original) data, where only minor differences are noted when compared to the mean and median sections. Visualization of uncertainty is represented by a map displaying the posterior standard deviation across realizations of the posterior estimates, depicted in Fig. 8(e). The high standard deviation is notable in the bottom left and right corners of the tomogram, primarily attributable to inadequate data coverage. The conductive topsoil displays low uncertainty, primarily due to its low resistivity and proximity to the surface electrodes. However, it is worth noting that the surface portions of the section beneath the hill area experience increased uncertainty due to topographical effects, and the model cells are not directly constrained by the measured

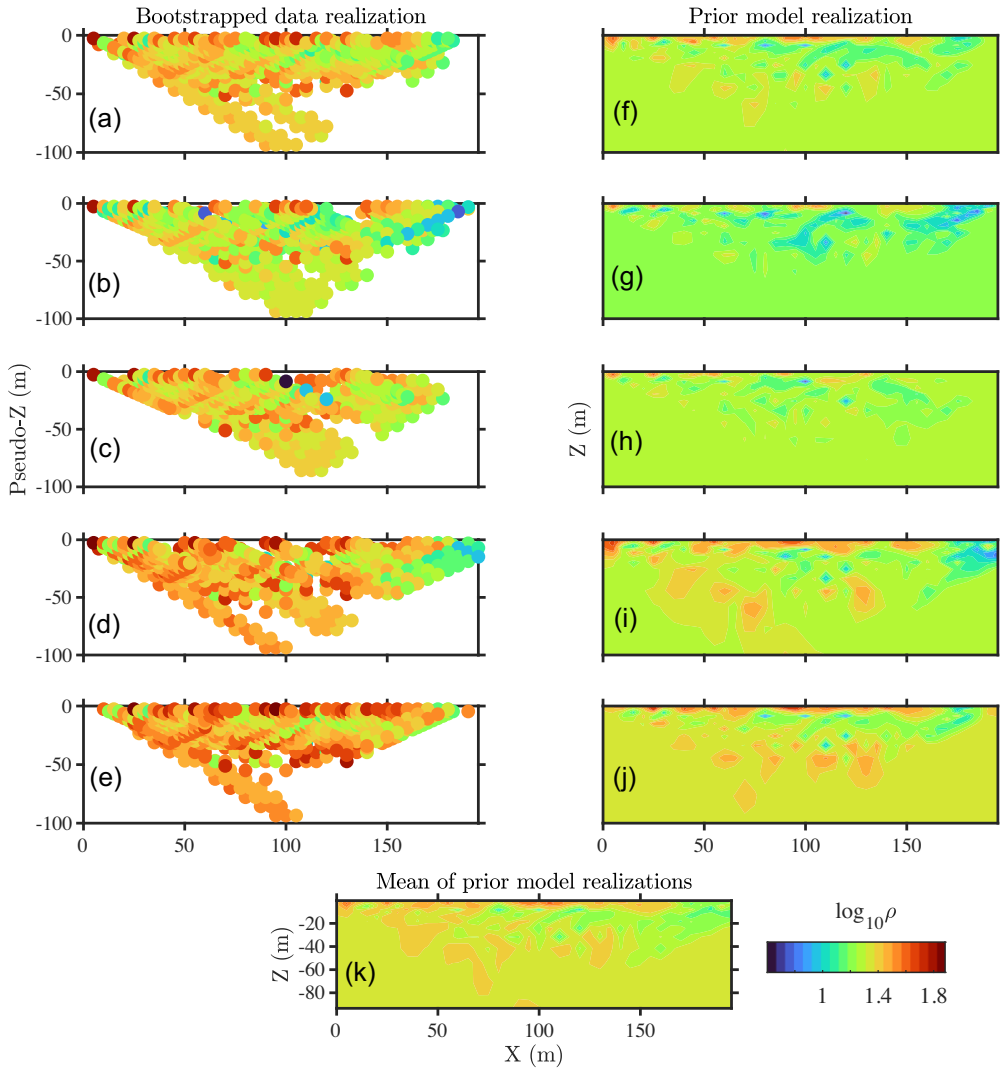

Figure 12. Example realizations of bootstrapped samples (a)–(e) with corresponding prior models generated using the geostatistical method (f)–(j), and the mean of prior model realizations (k) for real case 1.

data. Furthermore, the standard deviation plot illustrates an increase in uncertainty along the fault edge, showing that the precise slip of the fault cannot be determined.

In addition to visually analysing the uncertainty in the resistivity tomogram, histograms of the inverted resistivity values for the three locations represented by circle, square and triangle symbols are depicted in Fig. 9. In terms of the resolved parameters, the distributions show a bell shape with a single peak. However, some marginal distributions exhibit slight skewness either to the right or left.

Further insight can be gained through a comparison of the posterior resistivity distributions, the mean model and the true model associated with the three vertical profiles marked by black dashed lines (see Fig. 10). As illustrated in Fig. 10, a common feature of these plots is the increasing variability of posterior resistivity profiles in deeper sections, primarily attributed to reduced vertical resolution. Additionally, for transect P_2, the complexity of the fault zone results in a significant discrepancy between the model realizations and the ground truth. This discrepancy arises from the limitation of the ERT method in accurately recovering intricate geological structures.

3.2 Real examples

Finally, we demonstrate the application of the proposed method to two real data sets with known geological information. The first case focuses on the capability of the ERT method in detecting the soil–rock interface. In the second scenario, we address hydrogeological challenges by aiming to image the boundary between fresh and saline water. To enhance the resolution of sharp boundaries in both cases, we minimize the cost function (eq. 19) by setting $q = 1$.

Case 1

The first case study presented here is located near the Ilam embankment dam, within the western region of Iran. This study forms part of a comprehensive geophysical and geological investigation aimed at visualizing the boundary between topsoil and bedrock, and assessing the potential for a water escape zone on the left bank of the embankment dam. The primary challenge of this geophysical study involves the low conductivity contrast between clay and silt overburden layers and the limestone bedrock, which is interspersed with shale and marl formations. Out of the four parallel electrical resistivity profiles measured within the geophysical prospection

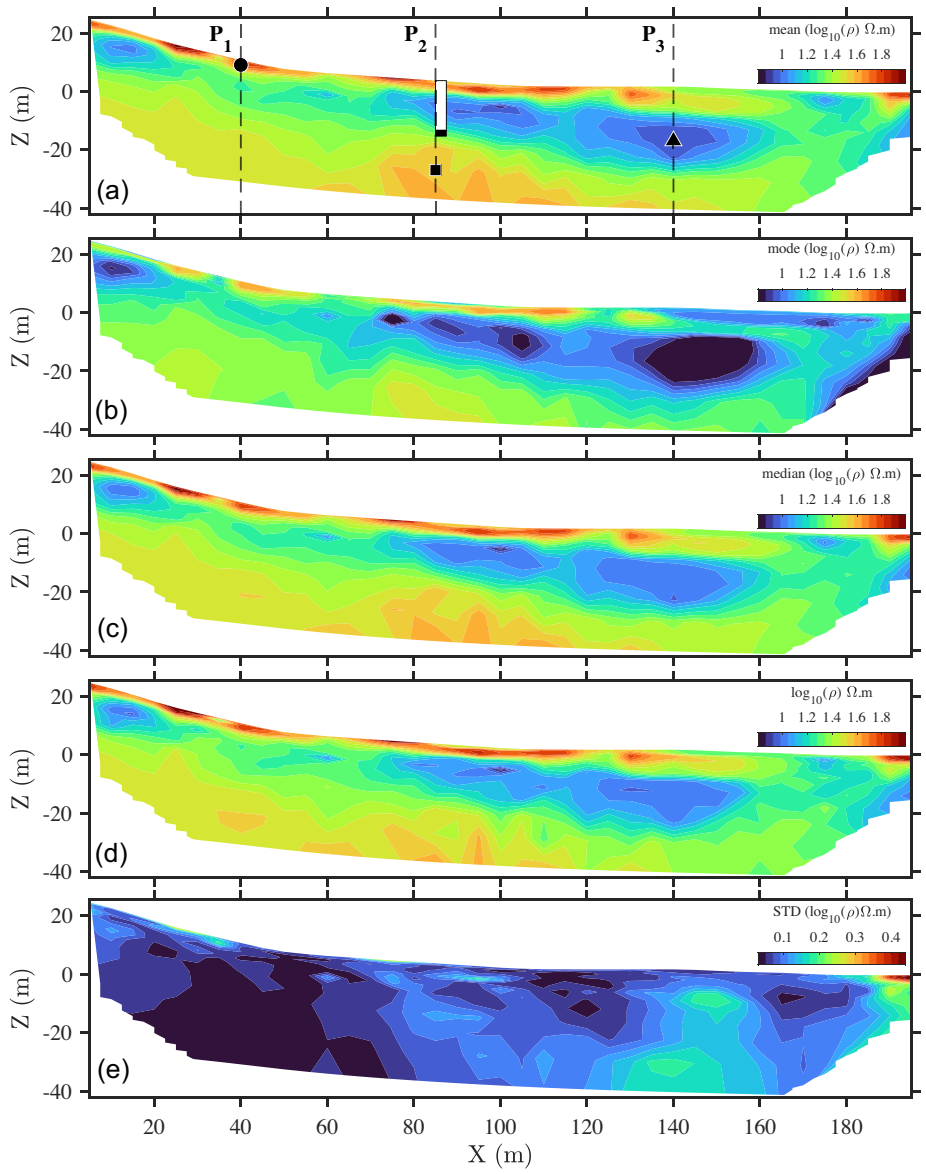

Figure 13. Statistical metrics derived from the inversion of real case 1 depicted in Fig. 11 using the proposed algorithm: (a) arithmetic mean, (b) mode (most frequent value), (c) median, (d) inverted section from the unbootstrapped (original) data and (e) standard deviation. Vertical dashed lines denote the location of the posterior resistivity profiles illustrated in Fig. 15. A lithology log acquired from borehole drilling overlays the mean section. White represents the silty clay overburden layer down to a depth of 15 m, while black signifies limestone bedrock interbedded with shales.

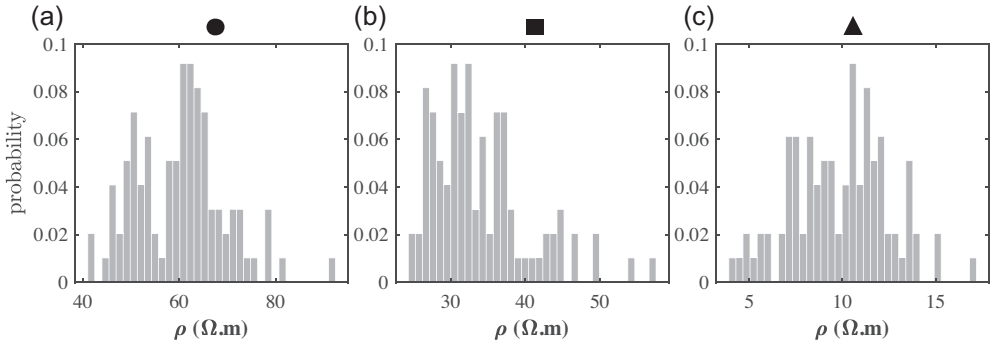

Figure 14. Posterior distributions of resistivity for the points depicted in Fig. 13, represented by (a) circle, (b) square and (c) triangle.

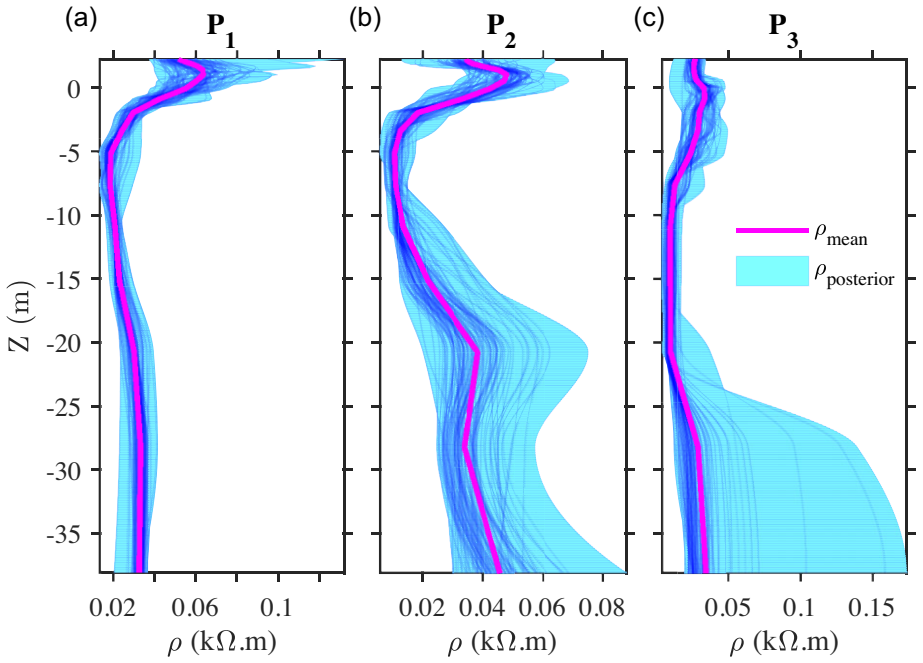

Figure 15. Vertical resistivity profiles along three transects (P_1, P_2 and P_3) shown in Fig. 13, obtained using inversion of the bootstrapped data. The shaded area indicates the inferred posterior resistivity distribution and the solid line represents the mean model.

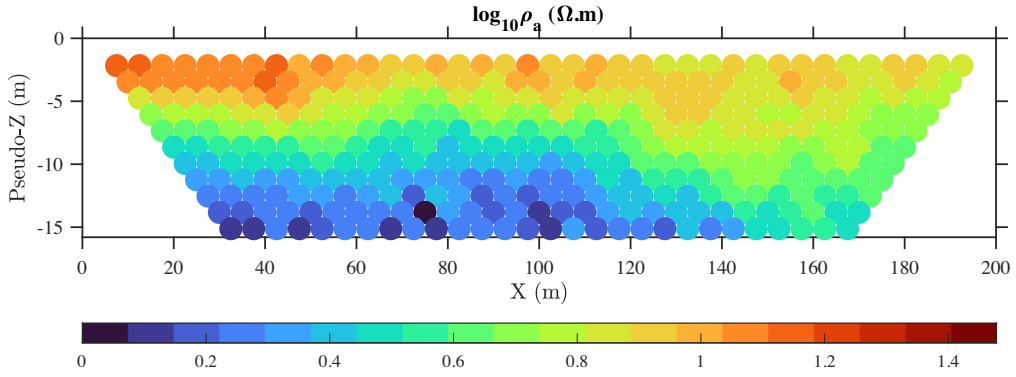

Figure 16. Resistivity pseudo-section associated with real case 2, acquired using a collinear Dipole–Dipole array with a minimum electrode spacing of 5 *m*.

area, one profile is considered in this case. The data acquisition was conducted using the eight-channel GDD instrument, together with a relay switching unit, along a profile comprising a total number of 959 data points. To achieve good vertical and horizontal resolution, a combination of the Dipole–Dipole and Wenner-alpha arrays was utilized with minimum electrode spacings of 5 and 10 *m* up to $s_{dd} = 8$ and $s_w = 12$ levels (where *s* is the number of current-potential dipole separation), respectively. The topography of the study area is characterized by a gently sloping surface. Nonetheless, we incorporate the topographical variations into the inversion process. According to the available lithology information obtained from the borehole in the vicinity of the ERT profile, there is a 15 *m*-thick overburden layer of silty clay above the limestone bedrock, which is interbedded with shales. However, the thickness of the overburden layer may exhibit slight variations along the profile line. To reduce the impact of noise on the inversion results, noisy data were filtered out of the field measurements. Data with an injected current of less than 10 *mA* or a standard deviation of electrical potential, derived from the stacking error, greater than 0.5 were discarded based on the filtering criterion. Fig. 11 illustrates the pseudo-plot representing the processed data with fewer data points for deeper parts of the pseudo-section. Upon qualitative analysis, despite the apparent resistivities increase with depth the demarcation between the topsoil and bedrock remains indistinct. In addition, resistivity values on the left side of the pseudo-section are generally greater than those on the right side.

Following the strategy outlined for synthetic examples, 100 realizations of perturbed data and prior models are created and used as inputs for inversion (see Algorithm 3), then the posterior distribution of the inverted resistivity models is obtained. To provide insights into the characteristics of the resampled data and the prior models generated through the geostatistical method, Fig. 12 showcases five randomly selected realizations of the bootstrapped data and their associated prior models.

As mentioned earlier, each realization is independently inverted from the others, allowing any number of CPUs to concurrently process data and initial model realizations without depending on one another. Thus this approach reduces the computing time of the algorithm implementation. Note that inversion parameters have to be identical across all bootstrapped data realizations to facilitate com-

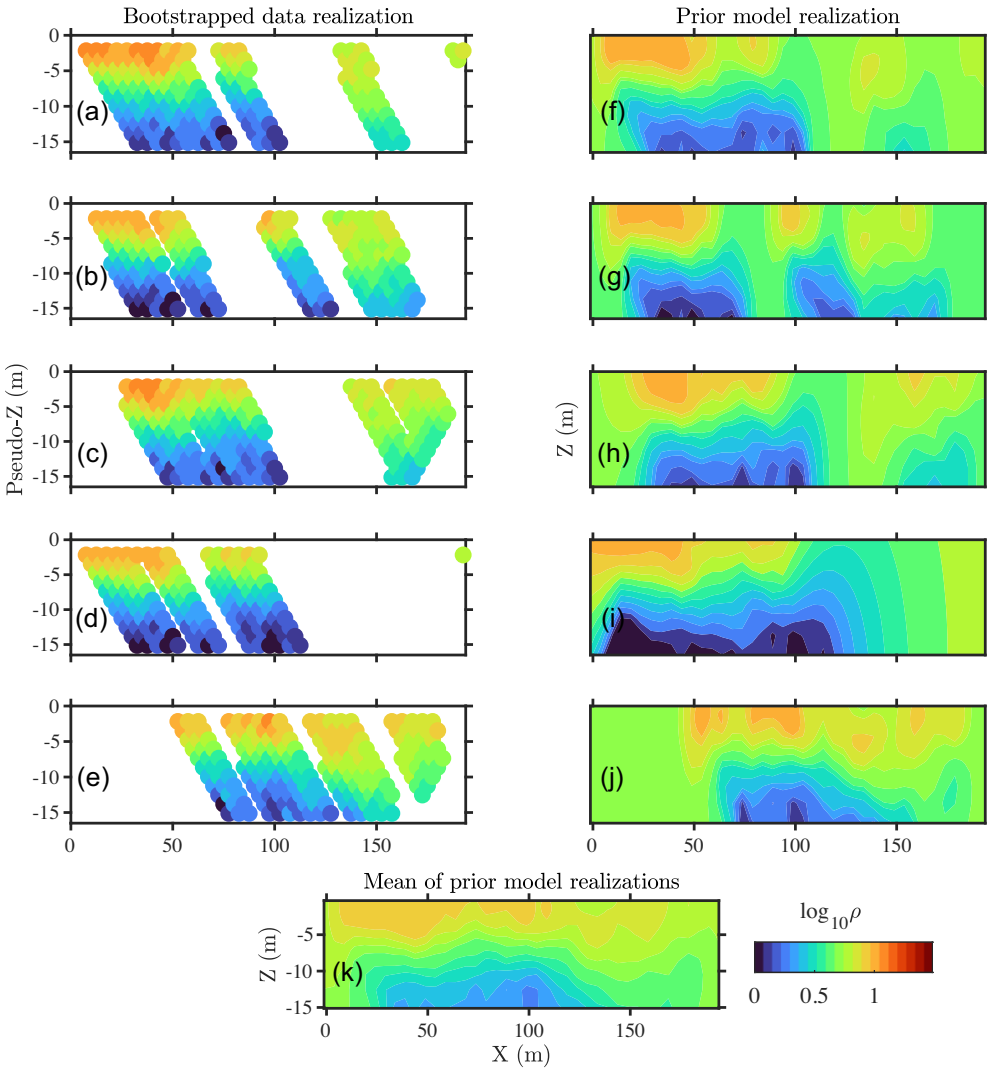

Figure 17. Example realizations of bootstrapped samples (a)–(e) with corresponding prior models generated using the geostatistical method (f)–(j), and the mean of prior model realizations (k) for the real case 2.

parability. However, in this manner, the uncertainty quantification will, to some degree, depend on the choice of starting parameter. After inverting the bootstrapped data realizations, the statistical properties of the posterior models including arithmetic mean, mode, median and standard deviation maps derived from convergent solutions of the proposed algorithm are illustrated in Figs 13(a)–(c). Just like in synthetic examples, convergence is achieved for each bootstrapped data realization after 10 iterations, depending on the number of continuation parameters ($\mathcal{T}$). The CPU time required for executing the entire inversion process is detailed in Table 2. The resistivity distribution displays horizontal layering, featuring a high-resistivity zone within the topmost two meters of the soil and a conductive middle layer with varying thickness along the profile. Additionally, the limestone bedrock is discernible as the most resistive layer at the bottom of the sections. When overlaying borehole information onto the posterior mean resistivity section, the interface between the topsoil and the bedrock aligns well with the borehole data. In the borehole lithology, white and black columns indicate the overburden layer and bedrock, respectively. In addition, to provide a comparison, Fig. 13(d) displays the inverted resistivity tomogram from the unbootstrapped (original) data, revealing only negligible differences when contrasted with the mean and median sections. Fig. 13(e) illustrates the posterior standard deviation, which mirrors the uncertainty of the inversion results. Generally, uncertainty escalates with depth, particularly in regions where model cells are minimally influenced by data, and primarily affected by regularization parameters. Notably, the high-density data set obtained through a combination of Wenner and Dipole–Dipole arrays results in the boundary between topsoil and bedrock not being depicted as a high-uncertainty zone.

Also, to assess resistivity recovery, we choose three points within the topmost layer (represented by a circle), bedrock (represented by a square) and overburden (represented by a triangle). Fig. 14 illustrates the posterior probability distribution of resistivity values associated with these model cells.

A more effective approach to scrutinizing potential uncertainties involves examining 1D posterior model realizations. We display 1D resistivity distributions along vertical profiles at points of interest in Fig. 15. From Fig. 15, it becomes apparent that the fluctuations in the inverted geo-electrical profiles generally intensify with depth, especially notable for the lowest layer (bedrock), attributed to reduced vertical resolution. In contrast to the depth of the resistive

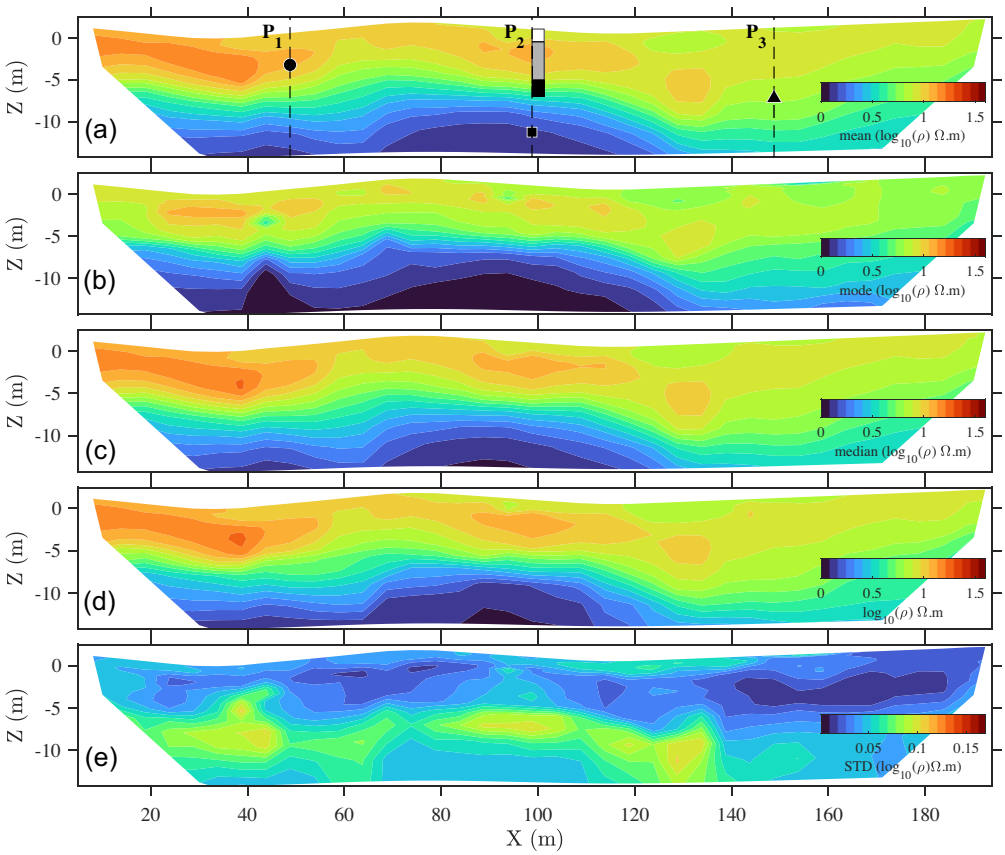

Figure 18. Statistical metrics derived from the inversion of the second real data presented in Fig. 16 using the proposed algorithm: (a) arithmetic mean, (b) mode (most frequent value), (c) median, (d) inverted section from the unbootstrapped (original) data and (e) standard deviation. Vertical dashed lines indicate the location of the posterior resistivity profiles shown in Fig. 20. A lithology log obtained from borehole drilling is superimposed on the mean section. In the lithology log, white denotes an unsaturated silt/clay zone down to a depth of 1.5 m, light grey indicates a saturated layer consisting of sand and clayey sand, and black points to the saline water-bearing layer.

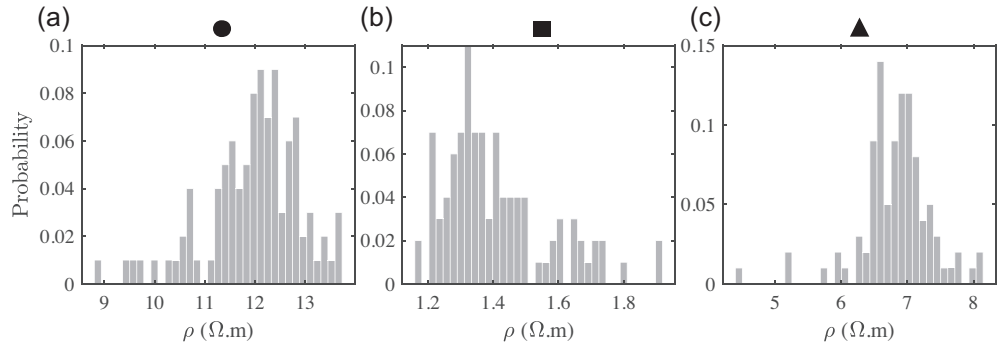

Figure 19. Posterior distributions of resistivity for the points depicted in Fig. 18, represented by (a) circle, (b) square and (c) triangle.

topmost layer, which lacks direct constraint by measured data, the estimation of the overburden-bedrock interface demonstrates high accuracy and low uncertainty, which is also compatible with the lithology log.

Case 2

The second set of real data was acquired in the Kashan Erg region, which is surrounded by sandy hills and salt pans. This area, located approximately 40 *km* from Kashan city in central Iran, boasts unique geographical features. Excessive groundwater withdrawal has caused a decline in the water table, leading to the intrusion of saline water into the freshwater aquifer. To address this issue, the fresh-saline water interface is resolved using ERT measurements, supplemented by a topographic survey to determine the profile's position. In this study, a collinear Dipole–Dipole configuration is employed, utilizing 23 electrodes arranged with a regular spacing of 5 m. This setup extends across eight levels ($s = 1 - 8$), resulting in 363 data points to generate the apparent resistivity responses. Although the Dipole–Dipole configuration offers benefits like less sensitivity to electromagnetic interference and minimal impact from measurement distortions caused by electrode polarization during current injection, this electrode setup often exhibits low-voltage

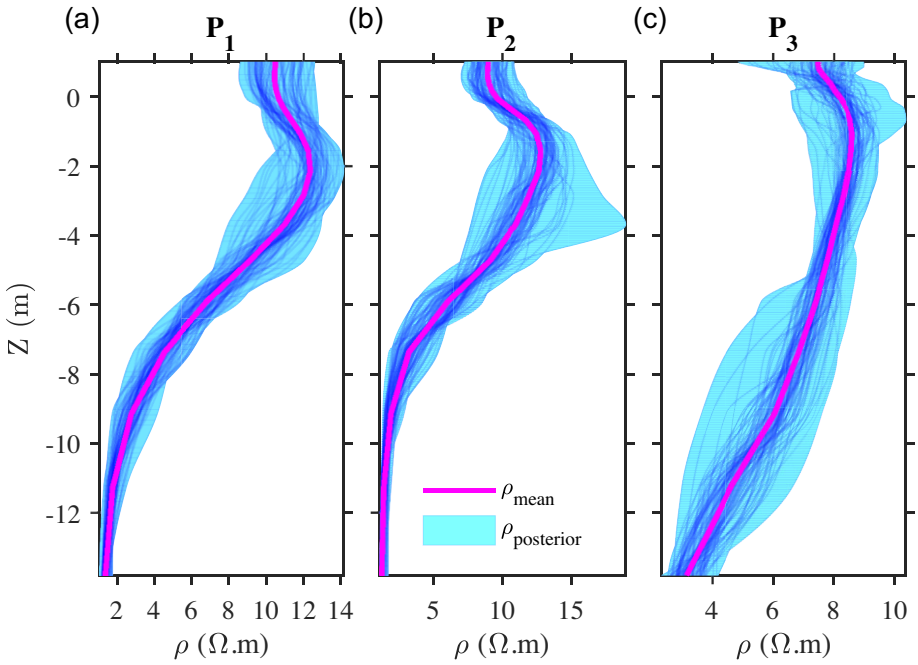

Figure 20. Vertical resistivity profiles along three transects (P_1, P_2 and P_3) shown in Fig. 18, obtained using inversion of the bootstrapped data. The shaded area indicates the inferred posterior resistivity distribution and the solid line represents the mean model.

strength when dealing with significant separations of current and potential dipoles. The data underwent filtering to eliminate outliers and measurements with low signal-to-noise ratios. The processed data are represented in Fig. 16 in the form of an apparent resistivity profile.

We adhere to the inversion strategy outlined in Algorithm 3, employing the proposed bootstrap resampling method to generate multiple data realizations. Subsequently, a range of prior models is stochastically generated for each bootstrapped data set. Following the implementation of the CBB algorithm and the geostatistical algorithm, Fig. 17 depicts five randomly selected realizations of the resampled data and their corresponding prior models. Displayed in Figs 18(a)–(c) are the descriptive statistics of the resulting solution ensemble, including the arithmetic mean, mode and median, respectively. The inverted sections reveal a heterogeneous distribution of subsurface resistivity, encompassing an unsaturated silt/clay zone and a saturated layer composed of sand and clayey sand. The saturated zone is distinctly divided into a freshwater-bearing layer with the highest resistivity values and a highly conductive layer containing saline water. The interface between saline and freshwater is identifiable at a depth of approximately 5 *m*. In addition, lithological information obtained from the available borehole is overlaid on the inversion results, revealing a relatively good agreement between the ground truth and the mean resistivity section. Likewise, the inversion of the unbootstrapped data is shown in Fig. 18(d) for comparison with the inverted tomograms of the resampled data. The difference between the unbootstrapped inversion and the mean and median sections is minimal. The standard deviation map, serving as a proxy for the uncertainty in the inverted subsurface model, is depicted in Fig. 18(e). Following this representation, it becomes apparent that the highest uncertainty lies in the fresh-saline water interface. However, the uncertainty decreases at the right hand of the section, likely due to the deeper penetration of the boundary between saline and fresh water. Additionally, plotting histograms of the reconstructed parameters offers a quick quantitative analysis.

Therefore, the distribution of inverted resistivity realizations corresponding to three points marked by a circle, square and triangle is represented in Fig. 19.

Furthermore, to evaluate resistivity recovery, the 1D posterior model realizations, along with the mean model associated with the three depth profiles depicted by black dashed lines in Fig. 18(a), are shown in Fig. 20. In these three plots, we depict the posterior resistivity distributions at different depths, represented in shaded colours, with the mean indicated by a solid magenta line. Upon examining the resistivity distributions of the first two transects (i.e. P_1 and P_2), it becomes evident that there is a greater uncertainty within the top 6 *m* of the profiles compared to the deeper depths. In the case of transect P_3, we observe an increase in the variations of resistivity distribution with depth, attributed to the presence of a more resistive region along the corresponding vertical profile. In this example, the elapsed CPU time required for the entire inversion is provided in Table 2.

4 CONCLUSIONS

Uncertainty quantification of electrical resistivity tomography models, derived from regularized inversion methods is an active field of research in geophysics. These models often carry inherent uncertainty, which can be attributed primarily to epistemic uncertainty, arising from imperfect underlying physics and inaccurate initial approximations of model parameters and aleatory variability, caused by measurement errors in observations. To address this issue, we propose a novel method in the framework of a stochastic optimization process for computing a meaningful uncertainty quantification for regularized inversion of ERT data. The proposed uncertainty quantification algorithm employs a hybrid approach, combining block-wise bootstrap resampling with a geostatistical method in the context of regularized inversion. To ensure that all data samples have an equal probability of being drawn during the bootstrapping

process, we utilized the CBB method. This method offers the advantage of wrapping samples around in a circular manner before blocking. Next, we apply a geostatistical technique to stochastically generate a set of initial models for each bootstrapped data set. Finally, a globally convergent homotopic continuation method is employed on each bootstrapped data set and initial model realization to explore the posterior resistivity models. We demonstrated the application of the proposed algorithm on two synthetic examples, inspired by real geological models, and on two real data sets with distinct regional geology. The objective was to retrieve statistical estimates of the resistivity distribution in the subsurface and to offer a quantitative analysis of the model uncertainty. From our numerical experiments, some general conclusions are drawn: (1) uncertainty tends to escalate in regions with fewer data constraints. (2) Areas exhibiting greater heterogeneity and complexity typically experience higher levels of uncertainty. (3) Uncertainty generally amplifies with depth, attributed to diminished vertical resolution.

In summary, the results showed that the proposed approach for quantifying uncertainty is straightforward to implement and easily integrates with any regularized inversion method within a specified set of (prior) assumptions. Practically speaking, the proposed method suggests that obtaining uncertainty estimates is achievable by executing the inversion algorithm in a parallel for-loop when sufficient resources are accessible. It is also noted that sampling the equivalence domain using the proposed method does not yield the desired Bayesian posterior distribution, however, from a practical standpoint, the estimated uncertainties hold significant value as they provide insights into the accuracy of the retrieved model features and enable the evaluation of result precision. The proposed uncertainty quantification method demonstrates computational efficiency suitable for addressing 2D ERT problems and can also be applied to 3D electrical resistivity inverse problems. However, one common limitation of the proposed method is the impact of prior assumptions, such as model parametrization and regularization term, on the posterior models and, consequently, on the resulting uncertainty quantification.

The inversion of ERT data inherently presents an imaging problem characterized by non-unique solutions. Addressing this through ensemble-based approaches, like our proposed algorithm, proves to be not only viable and useful but also poised to become increasingly practical in the future. As part of our future endeavor, we intend to conduct a comparative analysis between our approach and Bayesian inversion methods applied to ERT data.

CONFLICTS OF INTEREST

The authors declare that they have no financial interests or personal relationships that could have influenced the research presented in this paper.

ACKNOWLEDGMENTS

We would like to extend our sincere gratitude to the Editor-in-Chief, Alexis Maineult, as well as to Michael Tso and the anonymous reviewer for their interest in our work and their valuable contributions, which have greatly improved our manuscript.

DATA AVAILABILITY

The data associated with this research are available from the corresponding author upon reasonable request.

REFERENCES

Aleardi, M., Vinciguerra, A. & Hojat, A., 2021. A geostatistical Markov chain Monte Carlo inversion algorithm for electrical resistivity tomography, *Near Surf. Geophys.,* **19**(1), 7–26.

Aster, R., Borchers, B. & Thurber, C. H., 2005. *Parameter Estimation and Inverse Problems,* Elsevier Academic Press.

Axelsson, O. & Sysala, S., 2015. Continuation Newton methods, *Comput. Math. Appl.,* **70**(11), 2621–2637.

Beaujean, J., Nguyen, F., Kemna, A. & Engensgaard, P., 2010. Joint and sequential inversion of geophysical and hydrogeological data to characterize seawater intrusion models, in *Proceedings of the 21st Salt Water Intrusion Meeting,* pp. 21–25, Ponta Delgada, Portugal.

Binley, A. & Slater, L., 2020. *Resistivity and induced polarization: theory and applications to the near-surface earth,* Cambridge University Press, Cambridge

Blatter, D., Morzfeld, M., Key, K. & Constable, S., 2022. Uncertainty quantification for regularized inversion of electromagnetic geophysical data—Part I: motivation and theory, *Geophys. J. Int.,* **231**(2), 1057–1074.

Blatter, D., Morzfeld, M., Key, K. & Constable, S., 2022. Uncertainty quantification for regularized inversion of electromagnetic geophysical data–part II: application in 1-D and 2-D problems, *Geophys. J. Int.,* **231**(2), 1075–1095.

Bouchedda, A., Bernard, G. & Gloaguen, E., 2017. Constrained electrical resistivity tomography Bayesian inversion using inverse Matérn covariance matrix, *Geophysics,* **82**(3), E129–E141.

Campanya, J., Ledo, J., Queralt, P., Marcuello, A. & Jones, A.G., 2014. A new methodology to estimate magnetotelluric (MT) tensor relationships: Estimation of Local transfer-functions by Combining Interstation Transfer-functions (ELICIT), *Geophys. J. Int.,* **198**(1), 484–494.

Chilès, J.P. & Delfiner, P., 2012. *Geostatistics: Modeling Spatial Uncertainty,* Vol. **713,** John Wiley and Sons.

Christiansen, A.V. & Auken, E., 2012. A global measure for depth of investigation, *Geophysics,* **77**(4), WB171–WB177.

Chuang, J.M., Gui, Q.Y. & Hsiung, C.C., 1993. Numerical computation of Schwarz-Christoffel transformation for simply connected unbounded domain, *Comput. Methods Appl. Mech. Eng.,* **105**(1), 93–109.

Costamagna, E., 1987. On the numerical inversion of the Schwarz-Christoffel conformal transformation, *IEEE Trans. Microw. Theory Tech.,* **35**(1), 35–40.

Deceuster, J., Etienne, A., Robert, T., Nguyen, F. & Kaufmann, O., 2014. A modified DOI-based method to statistically estimate the depth of investigation of dc resistivity surveys, *J. Appl. Geophys.,* **103,** 172–185. doi: 10.1016/j.jappgeo.2014.01.018

de Pasquale, G., Linde, N. & Greenwood, A., 2019. Joint probabilistic inversion of DC resistivity and seismic refraction data applied to bedrock/regolith interface delineation, *J. Appl. Geophys.,* **170,** 103 839. doi: 10.1016/j.jappgeo.2019.103839

Dey, A. & Morrison, H.F., 1979. Resistivity modelling for arbitrarily shaped two-dimensional structures, *Geophys. Prospect.,* **27**(1), 106–136.

Dudek, A.E., 2015. Circular block bootstrap for coefficients of autocovariance function of almost periodically correlated time series, *Metrika,* **78**(3), 313–335.

Efron, B., 1979. Bootstrap methods: another look at the jackknife, *Ann. Statist.,* **7,** 1–26. doi: 10.1214/aos/1176344552

Efron, B. & Tibshirani, R.J., 1993. *An Introduction to the Bootstrap,* Chapman and Hall, New York.

Fernández-Muñiz, Z., Khaniani, H. & Fernández-Martínez, J.L., 2019. Data kit inversion and uncertainty analysis, *J. Appl. Geophys.,* **161,** 228–238. doi: 10.1016/j.jappgeo.2018.12.022

Friedel, S., 2003. Resolution, stability and efficiency of resistivity tomography estimated from a generalized inverse approach, *Geophys. J. Int.,* **153**(2), 305–316.

Galetti, E. & Curtis, A., 2018. Transdimensional electrical resistivity tomography, *J. Geophys. Res.: Solid Earth,* **123**(8), 6347–6377.

Ghanati, R. & Müller-Petke, M., 2021. A homotopy continuation inversion of geoelectrical sounding data, *J. Appl. Geophys.,* **191,** 104 356. doi: 10.1016/j.jappgeo.2021.104356

Ghanati, R. & Fallahsafari, M., 2022. Fréchet Derivatives calculation for electrical resistivity imaging using forward matrix method, *Iran. J. Geophys.,* **15**(4), 153–163.

Ghanati, R. & Fallahsafari, M., 2023. Incorporating topographic variations on electrical resistance tomography, *Int. J. Min. Geo-Eng.,* **57**(3), 335–341.

Gouveia, W.P. & Scales, J.A., 1997. Resolution of seismic waveform inversion: Bayes versus Occam, *Inverse Probl.,* **13**(2), 323.

Grayver, A.V. & Kuvshinov, A.V., 2016. Exploring equivalence domain in nonlinear inverse problems using Covariance Matrix Adaption Evolution Strategy (CMAES) and random sampling, *Geophys. J. Int.,* **205**(2), 971–987.

Han, B., Fu, H.S. & Li, Z., 2005. A homotopy method for the inversion of a two-dimensional acoustic wave equation, *Inverse Probl. Sci. Eng.,* **13**(4), 411–431.

Hertrich, M., 2008. Imaging of groundwater with nuclear magnetic resonance, *Prog. Nucl. Magn. Reson. Spectrosc.,* **53**(4), 227–248.

Kaipio, J. & Somersalo, E., 2007. Statistical inverse problems: discretization, model reduction and inverse crimes, *J. Comput. Appl. Math.,* **198**(2), 493–504.

Kalscheuer, T., De los Ángeles García Juanatey, M., Meqbel, N. & Pedersen, L.B., 2010. Non-linear model error and resolution properties from two-dimensional single and joint inversions of direct current resistivity and radiomagnetotelluric data, *Geophys. J. Int.,* **182**(3), 1174–1188.

Kemna, A., 2000. Tomographic inversion of complex resistivity, *Ph.D. Thesis*, *Ruhr-Universität Bochum,* pp. **169**

Künsch, H.R., 1989. The jackknife and the bootstrap for general stationary observations, *Ann. Statist.,* **17**(3), 1217–1241.

Lahiri, S.N., 1999. Theoretical comparisons of block bootstrap methods, *Ann. Statist.,* **27**(1), 386–404.

Liao, S., 2004. On the homotopy analysis method for nonlinear problems, *Appl. Math. Comput.,* **147**(2), 499–513.

Liu, R.Y. & Singh, K., 1992. Moving blocks jackknife and bootstrap capture weak dependence, In: Lepage, R. & Billard, L.eds., *Exploring the Limits of Bootstrap,* Wiley, New York

Matheron, G., 1971. Les Cahiers du Centre de Morphologie Mathématique, *The Theory of Regionalised Variables and Its Applications,* **Vol. 5,** pp. 212, Ecole Nationale Superieure des Mines de Paris

McLaughlin, K.L., 1988. Maximum-likelihood event magnitude estimation with bootstrapping for uncertainty estimation, *Bull. seism. Soc. Am.,* **78**(2), 855–862.

Mosegaard, K. & Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems, *J. Geophys. Res.: Solid Earth,* **100**(B7), 12431–12447.

Neukirch, M. & García, X., 2014. Nonstationary magnetotelluric data processing with instantaneous parameter, *J. Geophys. Res.: Solid Earth,* **119**(3), 1634–1654.

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

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

Oldenborger, G.A., Routh, P.S. & Knoll, M.D., 2007. Model reliability for 3D electrical resistivity tomography: Application of the volume of investigation index to a time-lapse monitoring experiment, *Geophysics,* **72**(4), F167–F175.

Oldenborger, G.A. & Routh, P.S., 2009. The point-spread function measure of resolution for the 3-D electrical resistivity experiment, *Geophys. J. Int.,* **176**(2), 405–414.

Parsekian, A.D. & Grombacher, D., 2015. Uncertainty estimates for surface nuclear magnetic resonance water content and relaxation time profiles from bootstrap statistics, *J. Appl. Geophys.,* **119,** 61–70. doi: 10.1016/j.jappgeo.2015.05.005

Patton, A., Politis, D.N. & White, H., 2009. Automatic Block-Length Selection for the Dependent Bootstrap, *Econom. Rev.,* **28**(4), 372–375.

Politis, D.N. & Romano, J.P., 1992. A circular block-resampling procedure for stationary data, *Exploring the Limits of Bootstrap. Wiley Series in Probability and Mathematical Statistics,* Wiley, New York.

Politis, D.N. & Romano, J.P., 1995. Bias-corrected nonparametric spectral estimation, *J. Time Ser. Anal.,* **16**(1), 67–103.

Politis, D.N. & White, H., 2004. Automatic block-length selection for the dependent bootstrap, *Econom. Rev.,* **23**(1), 53–70.

Pryet, A., Ramm, J., Chilès, J.P., Auken, E., Deffontaines, B. & Violette, S., 2011. 3D resistivity gridding of large AEM datasets: A step toward enhanced geological interpretation, *J. Appl. Geophys.,* **75**(2), 277–283.

Ramirez, A.L., Daily, W.D. & Newmark, R.L., 1995. Electrical resistance tomography for steam injection monitoring and process control, *J. Environ. Eng. Geophys.,* **1**(A), 39–51.

Ramirez, A.L. *et al.*, 2005. Stochastic inversion of electrical resistivity changes using a Markov Chain Monte Carlo approach, *J. Geophys. Res.: Solid Earth,* **110**(B2). doi: 10.1029/2004JB003449

Rion, W.L. & Van Brunt, V., 1990. Differential geometry based homotopy continuation, *Comput. Chem. Eng.,* **14**(8), 889–900.

Roudsari, M. S., Ghanati, R. & Bérubé, C. L., 2024. Spectral induced polarization tomography inversion: Hybridizing homotopic continuation with Bayesian inversion, *Geophysics,* **89**(4), 1–63.

Rosas-Carbajal, M., Linde, N., Kalscheuer, T. & Vrugt, J.A., 2014. Two-dimensional probabilistic inversion of plane-wave electromagnetic data: methodology, model constraints and joint inversion with electrical resistivity data, *Geophys. J. Int.,* **196**(3), 1508–1524.

Sarma, D.D., 2010. *Geostatistics with Applications in Earth Sciences,* Springer Science and Business Media.

Schnaidt, S. & Heinson, G., 2015. Bootstrap resampling as a tool for uncertainty analysis in 2-D magnetotelluric inversion modelling, *Geophys. J. Int.,* **203**(1), 92–106.

Shearer, P.M., 1997. Improving local earthquake locations using the L1 norm and waveform cross correlation: Application to the Whittier Narrows, California, aftershock sequence, *J. Geophys. Res.: Solid Earth,* **102**(B4), 8269–8283.

Slater, L., Binley, A.M., Daily, W. & Johnson, R., 2000. Cross-hole electrical imaging of a controlled saline tracer injection, *J. Appl. Geophys.,* **44**(2–3), 85–102.

Stein, M.L., 1999. *Interpolation of Spatial Data: Some Theory for Kriging,* Springer Science and Business Media.

Tichelaar, B.W. & Ruff, L.J., 1989. How good are our best models? Jackknifing, bootstrapping, and earthquake depth, *Eos, Trans. Am. Geophys. Union,* **70**(20), 593–606.

Tso, C.H.M. *et al.*, 2017. Improved characterisation and modelling of measurement errors in electrical resistivity tomography (ERT) surveys, *J. Appl. Geophys.,* **146,** 103–119. doi: 10.1016/j.jappgeo.2017.09.009

Tso, C.H.M., Iglesias, M., Wilkinson, P., Kuras, O., Chambers, J. & Binley, A., 2021. Efficient multiscale imaging of subsurface resistivity with uncertainty quantification using ensemble Kalman inversion, *Geophys. J. Int.,* **225**(2), 887–905.

Tso, C.H.M., Iglesias, M. & Binley, A., 2024. Ensemble Kalman inversion of induced polarization data, *Geophys. J. Int.,* **236**(3), 1877–1900.

Valavi, R., Elith, J., Lahoz-Monfort, J.J. & Guillera-Arroita, G., 2019. blockCV: An R package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution models, *Methods Ecol. Evol.,* **10**(2), 225–232.

Vinciguerra, A., Aleardi, M., Hojat, A., Loke, M.H. & Stucchi, E., 2021. Discrete cosine transform for parameter space reduction in Bayesian electrical resistivity tomography, *Geophys. Prospect.,* **70**(1), 193–209.

Wackernagel, H., 2003. *Multivariate Geostatistics,* 3rd edn. Springer-Verlag, Berlin. doi: 10.1007/978-3-662-05294-5

Watson, L.T. & Haftka, R.T., 1989. Modern homotopy methods in optimization, *Comput. Methods Appl. Mech. Eng.,* **74**(3), 289–305.

Wei, X. & Sun, J., 2021. Uncertainty analysis of 3D potential-field deterministic inversion using mixed Lp norms, *Geophysics,* **86**(6), G133–G158.

Yang, X., Chen, X., Carrigan, C.R. & Ramirez, A.L., 2014. Uncertainty quantification of CO2 saturation estimated from electrical resistance tomography data at the Cranfield site, *Int. J. Greenhouse Gas Control,* **27,** 59–68. doi: 10.1016/j.ijggc.2014.05.006