Journal of Applied Geophysics 196 (2022) 104512

Contents lists available at ScienceDirect

Journal of Applied Geophysics

journal homepage: www.elsevier.com/locate/jappgeo

A comprehensive workflow to analyze ensembles of globally inverted 2D electrical resistivity models

Mauricio Arboleda-Zapata *, Julien Guillemoteau, Jens Tronicke

Universität Potsdam, Institut für Geowissenschaften, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

ARTICLE INFO

Keywords:
Near-surface geophysics
Electrical resistivity tomography
Non-uniqueness
Global inversion
Particle swarm optimization
Ensemble analysis

ABSTRACT

Electrical resistivity tomography (ERT) aims at imaging the subsurface resistivity distribution and provides valuable information for different geological, engineering, and hydrological applications. To obtain a subsurface resistivity model from measured apparent resistivities, stochastic or deterministic inversion procedures may be employed. Typically, the inversion of ERT data results in non-unique solutions; i.e., an ensemble of different models explains the measured data equally well. In this study, we perform inference analysis of model ensembles generated using a well-established global inversion approach to assess uncertainties related to the non-uniqueness of the inverse problem. Our interpretation strategy starts by establishing model selection criteria based on different statistical descriptors calculated from the data residuals. Then, we perform cluster analysis considering the inverted resistivity models and the corresponding data residuals. Finally, we evaluate model uncertainties and residual distributions for each cluster. To illustrate the potential of our approach, we use a particle swarm optimization (PSO) algorithm to obtain an ensemble of 2D layer-based resistivity models from a synthetic data example and a field data set collected in Loon-Plage, France. Our strategy performs well for both synthetic and field data and allows us to extract different plausible model scenarios with their associated uncertainties and data residual distributions. Although we demonstrate our workflow using 2D ERT data and a PSO-based inversion approach, the proposed strategy is general and can be adapted to analyze model ensembles generated from other kinds of geophysical data and using different global inversion approaches.

1. Introduction

Electrical resistivity tomography (ERT) is a popular geophysical method typically used in different hydrological, geotechnical, environmental, geological, and engineering applications (Loke et al., 2013). In resistivity surveying, the standard workflow consists of data acquisition, processing, inversion, and interpretation. During 2D ERT data acquisition, two electrodes inject an electrical current into the ground while another pair of electrodes measure the voltage. Such quadripole configurations are performed at different positions and spacings along a profile to obtain a 2D pseudo-section of the apparent resistivities calculated from the measured voltages. The next step is data evaluation and processing, which typically consists of checking data quality and editing/filtering erroneous readings (e.g., related to electrode coupling problems). Subsequently, the filtered data are inverted to reconstruct a 2D resistivity model of the subsurface. Finally, inverted resistivity models are interpreted (e.g., in terms of geological structures) typically considering the sensitivity matrix (e.g., Oldenburg and Li, 1999; Furman et al., 2003) to exclude unresolved model areas from the interpretation.

According to Günther et al. (2006) and Loke et al. (2013), if the data set is incomplete, inconsistent, or noisy, the inverse problem is ill-posed and characterized by a non-unique solution; i.e., an ensemble of model solutions describes the data equally well (e.g., Fernández-Martínez et al., 2017; De Pasquale et al., 2019). Different approaches have been suggested to generate an ensemble of models that are consistent with a given geophysical data set. On the one hand, there are strategies that use local inversion routines considering different initial conditions (e.g., Vasco et al., 1996) and/or regularization parameters (e.g., Constable et al., 2015). On the other hand, there are strategies whose objective is to obtain solutions independent of starting models or regularization parameters using global optimization approaches. For example, such global inversion strategies are based on genetic algorithms (e.g., Stoffa and Sen, 1991; Akça and Basokur, 2010), particle swarm optimization (e.g., Shaw and Srivastava, 2007; Fernández-Martínez et al., 2010; Tronicke et al., 2012), simulated annealing (e.g., Rothman, 1985; Sen and Stoffa, 1991), and Monte Carlo methods (e.g., Tarantola, 2005;

* Corresponding author.
E-mail address: arboleda@geo.uni-potsdam.de (M. Arboleda-Zapata).

https://doi.org/10.1016/j.jappgeo.2021.104512
Received 25 March 2021; Received in revised form 15 November 2021; Accepted 25 November 2021
Available online 29 November 2021

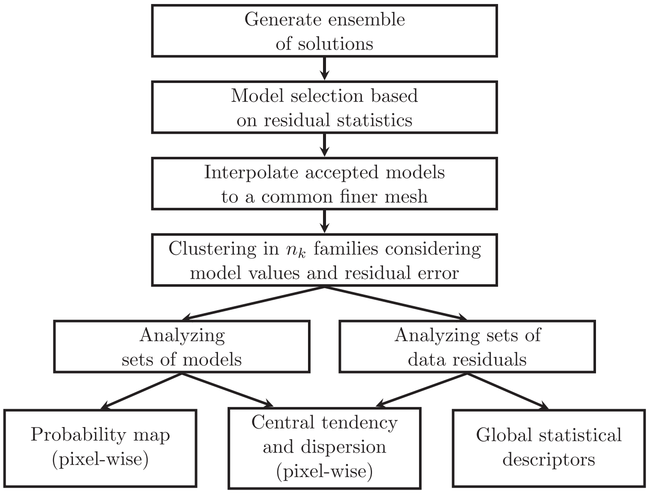

Fig. 1. Flow diagram illustrating the proposed interpretation strategy of model ensembles containing potential solutions of a 2D ERT inversion problem. The individual steps of this workflow are explained in more detail in Sections 2.1 to 2.6.

Cordua et al., 2012).

To analyze and interpret the generated ensemble of solutions, different approaches have been proposed in the literature. For example, Sen and Stoffa (1996) and de Groot-Hedlin and Vernon (1998) analyzed the correlation matrix to estimate the trade-off between individual model parameters. Sambridge (1999) used a neighborhood algorithm to resample the parameter space and, then, calculated the resolution matrix and the trade-off in the model parameter space. Tompkins et al. (2011) proposed a strategy called cut-off maps which consists of a pixel-wise calculation of specific quantiles of the model ensemble. Tronicke et al. (2012), Fernández-Martínez et al. (2012), and Zhu and Gibson (2018) used statistical measures such as the ensemble mean or median to characterize the central tendency of the ensemble whose uncertainties are assessed through dispersion measures such as the standard deviation or the interquartile range. Furthermore, Hermans et al. (2015), Fernández-Martínez et al. (2017), and De Pasquale et al. (2019) applied probability maps to study the uncertainties for a particular range of model parameters such as resistivities associated, for example, with a specific geologic unit, and Scheidt et al. (2018) studied the variability of a set of models using orthogonal component analysis.

Whether to use the entire ensemble of solutions for posterior analysis or not, is a question that has been addressed in previous studies (e.g., Lomax and Snieder, 1995; Douma et al., 1996; Vasco et al., 1996; Sambridge, 1999; Douma and Haney, 2013). Lomax and Snieder (1995) in a synthetic experiment of seismic dispersion curves apply a threshold criteria accepting the inverted models which are less than 0.85 times the applied noise level of the synthetic data to insure suitable model solutions. Douma et al. (1996) and Douma and Haney (2013) used empirical orthogonal functions to reject the models that add the most variability to the ensemble. Alternatively, Vasco et al. (1996) and Sambridge (1999) consider all model solutions but introduce different weighting criteria to derive posterior statistics. Such an approach helps to consider information that may be contained in not converged models and, thus, might also be useful to guide resampling strategies (Sambridge, 1999). Barboza et al. (2018, 2019) proposed a quality measure for a set of solutions considering both the global error and the quantile 90% of the residuals (i.e., the differences between the measured and calculated data). Vasco et al. (1996), Ramirez et al. (2005), and Fernández-Martínez et al. (2017) used cluster analysis strategies to group models that are similar to each other and, then, performed posterior statistical analysis for each cluster.

In this study, we adapt some of these ideas and propose a logical and unified workflow to interpret ensembles of 2D ERT inversion results. Our ensemble inference strategy starts with formulating a model rejection criterion that considers different statistical measures calculated from the residuals. Then, we evaluate the variability of the resulting reduced ensemble, including the identification of potential model groupings using cluster analysis. Finally, we assess model uncertainties including the structure of residuals for each identified group of models. To illustrate our workflow, we use a global inversion strategy for generating an ensemble of possible model solutions for two 2D ERT data sets. The first data set corresponds to a synthetic example created considering a three-layer input model, while the second one represents a field data set collected in a coastal environment in Loon-Plage, France.

2. Methodology

We summarize our workflow for ensemble interpretation in Fig. 1. In the first step, we create an ensemble of resistivity models representing possible solutions for the considered 2D ERT inverse problem. Then, we filter this ensemble considering a statistical analysis of the corresponding residual distributions. Previous to any further posterior analysis, we interpolate all models to a common mesh to allow for pixel-wise calculations and analyses. Next, we perform cluster analysis to group the models and their corresponding residual data. Finally, we use the models from each cluster to assess uncertainties using probability maps and statistical descriptors of the central tendency and dispersion. Similarly, we use the data residuals from each cluster to evaluate the residual distribution in a pixel-wise fashion using global statistical descriptors. In the following subsections, we explain each of these steps in more detail.

2.1. Generating an ensemble of solutions

To generate an ensemble consisting of n_m solutions (the first step in Fig. 1), we invert our 2D ERT data sets using a particle swarm optimization (PSO) algorithm considering a layer-based model parameterization as described in more detail by Tronicke et al. (2012) and Rumpf and Tronicke (2014, 2015). Our model parameterization relies on sums of arc-tangent functions (Roy et al., 2005), which represent the interfaces between individual layers. For a single interface, this can be written as

$$z(x) = z_0 + \sum_{j=0}^{n_{nod}} \Delta z_j \left(0.5 + \frac{1}{\pi} tan^{-1} \left(x - \frac{x_j}{b_j} \right) \right), \tag{1}$$

where z is the depth, n_{nod} is the number of arc-tangent nodes, z_0 is the average depth of the interface, x_j is the horizontal location of an arc-tangent node, and Δz_j is the vertical throw attained asymptotically over a horizontal distance of b_j. One advantage of this strategy is its flexibility because Eq. (1) allows the representation of various interface geometries including, for example, step-like or smooth variations with depth. When inverting 2D ERT data using this model parameterization, the number of model parameters n_{par} required to describe a resistivity model $\mathbf{m}$ consisting of n_{int} interfaces is given by

$$n_{par} = \left(n_{int} + 3 \sum_{i=1}^{n_{int}} n_{nod_i} \right) + (n_{int} + 1), \tag{2}$$

where the first term indicates the number of parameters needed to describe the interface structure and the second term the number of homogeneous layers (i.e., layers with specific resistivity ρ). It should be noted that such a layer-based parameterization may introduce some bias, and some features may not be resolved (Rosas-Carbajal et al., 2014). However, if a priori information (e.g., regarding the geological settings) indicates a layered subsurface, such an approach may help to provide more realistic model solutions (e.g., Smith et al., 1999; Auken and Christiansen, 2004; Boiero and Socco, 2014; Juhojuntti and Kamm, 2015; De Pasquale et al., 2019) and to reduce the number of model parameters making it attractive for global inversion approaches such as the PSO-based strategy used in this study.

During PSO-based inversion, a swarm of particles explores the model-parameter space and the goodness of a model (i.e., parameter combination) is evaluated by a predefined objective function. Here, we use the root mean square logarithmic error *RMSLE* defined as

$$RMSLE = \sqrt{\frac{1}{n_d}\sum_{i=0}^{n_d} e^2}, \tag{3}$$

where n_d is the number of data points, and $\mathbf{e}$ is the vector of residuals calculated by

$$\boldsymbol{e} = log\left(\boldsymbol{d}^{cal}\right) - log\left(\boldsymbol{d}^{obs}\right). \tag{4}$$

Here, $\boldsymbol{d}^{obs}$ represents the observed and $\boldsymbol{d}^{cal}$ the calculated data vector (note that $\boldsymbol{e}$ can be rewritten as $log(\boldsymbol{d}^{cal}/\boldsymbol{d}^{obs})$ which results in a dimensionless quantity). To obtain $\boldsymbol{d}^{cal}$, we use the forward operator as implemented in the software package BERT (Günther et al., 2006; Rücker et al., 2006) considering an adaptive unstructured mesh (i.e., a new mesh is generated for each particle and for each iteration during the PSO-based inversion) that is created using the python library pyGIMLi (Rücker et al., 2017). During optimization, the PSO algorithm assesses two points of attraction; i.e., the best solution of each individual particle and the best solution of all particles. Together with some random perturbations, these control the movement of the particle swarm in the model parameter space. If the optimization meets a predefined stopping criterion, such as a threshold of the objective function or a maximum number of iterations, we save the final best model. By repeating the inversion using different seeds of the random number generator, we generate an ensemble of independent model solutions $\boldsymbol{M} = [\boldsymbol{m}_1, \boldsymbol{m}_2, \ldots, \boldsymbol{m}_{n_m}]$ and the corresponding residuals $\boldsymbol{\delta} = [\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_{n_m}]$. Further details regarding this PSO-based global inversion strategy are given by Fernández-Martínez et al. (2010) and Tronicke et al. (2012). To evaluate if we obtained a representative ensemble of models, we inspect that there are no significant variations in the posterior correlation matrix after adding more model solutions to our ensemble as outlined by Sen and Stoffa (2013).

2.2. Model selection process

Although we employ a global inversion method, some optimization runs might not converge; i.e., get trapped in a local minimum and resulting in an inverted model with poorly fitted data. Considering these models in the posterior analysis could be misleading for the interpretation (e.g., Lomax and Snieder, 1995; Douma et al., 1996). Thus, we use a model selection criterion based on globally evaluating $\boldsymbol{\delta}$ (the second step in Fig. 1) considering the root mean square logarithmic error $RMSLE(\boldsymbol{\delta})$, the interquartile range $IQR(\boldsymbol{\delta})$, and the quantile 90% $q_{90}(\boldsymbol{\delta})$. In the end, we accept models satisfying $RMSLE(\boldsymbol{\delta}) \leq \mathrm{t}_1$, $IQR(\boldsymbol{\delta}) \leq \mathrm{t}_2$, and $q_{90}(\boldsymbol{\delta}) \leq \mathrm{t}_3$, where t_1, t_2, and t_3 are thresholds defined empirically or by considering error estimates from repetition or reciprocity measurements (e.g., Zhou and Dahlin, 2003). However, in our data sets such information is not available and, thus, we use histograms and scatter plots of $RMSLE(\boldsymbol{\delta})$, $IQR(\boldsymbol{\delta})$, and $q_{90}(\boldsymbol{\delta})$ as tools to estimate plausible thresholds values to remove those models and residuals from $\boldsymbol{M}$ and $\boldsymbol{\delta}$, respectively. A further complementary visualization strategy (not implemented in this study) could be to plot the cumulative distribution function which might be helpful to filter by a quantile criteria (e.g., accepting 80% of the models in $\boldsymbol{M}$). These simple but informative visualization strategies may need some user validation, for example, by inspecting the models around the defined threshold. Once we are confident with the chosen threshold, we define a reduced set of accepted models $\boldsymbol{M}_{acc}$ and residuals $\boldsymbol{\delta}_{acc}$

2.3. Interpolation to a common mesh

Our layer-based model parameterization strategy results in a different finite-element mesh for each model in $\boldsymbol{M}_{acc}$ (as needed by the used forward solver). To perform posterior statistical analyses, these models are interpolated (the third step in Fig. 1) using the nearest-neighbor technique on a densely discretized mesh; i.e., to ensure no critical loss of model resolution. Having all models with the same discretization allows us to perform pixel-wise calculations and analyses to derive, for example, the central trend and dispersion models.

2.4. Cluster analysis

We perform cluster analysis to identify and separate possible groups of models reflecting different subsurface scenarios. The most straightforward way to cluster the resulting model ensemble is to define the cluster input matrix as $\boldsymbol{K} = \boldsymbol{M}_{acc}$ (e.g., Fernández-Martínez et al., 2017). Alternatively, one could also set $\boldsymbol{K} = \boldsymbol{\delta}_{acc}$ in order to group models that show similar misfit distributions and patterns within the pseudo-section. Although these two strategies may provide reasonable results independently, in this study, we propose that defining $\boldsymbol{K}$ considering information from both $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$ might potentially improve the clustering output (the fourth step in Fig. 1).

When clustering using $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$ together, we have to consider resampling and rescaling (because $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$ have different sizes and physical meanings) to ensure a well balanced influence from $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$ on the result of cluster analysis. Typically, each model vector in $\boldsymbol{M}_{acc}$ is significantly larger than the number of corresponding data residuals. Therefore, if we consider $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$ with their original sizes, the cluster output might be biased towards $\boldsymbol{M}_{acc}$. To avoid such a bias, we first resample the models in $\boldsymbol{M}_{acc}$ in to a set of vertical resistivity profiles to obtain comparable sizes of $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$. Note that for a grid-based parameterization strategy, random resampling may be more appropriated than resampling along vertical profiles, also other resampling strategies might be considered, but they are out of the scope of this study. Furthermore, because $\boldsymbol{M}_{acc}$ values are given in Ωm and $\boldsymbol{\delta}_{acc}$ dimensionless, we standardize $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$ such that they show a mean of zero and standard deviation of one. With this, our input matrix $\mathbf{K}$ is defined as the concatenation

$$\boldsymbol{K} = \boldsymbol{M}_{acc}^{r,s} \frown \boldsymbol{\delta}_{acc}^{s}, \tag{5}$$

where the superscripts r and s indicate resampled and scaled variables, respectively.

To perform a cluster analysis of $\mathbf{K}$, we use the well-known k-means algorithm (MacQueen, 1967). In the k-means algorithm, we must predefine the number of clusters n_k. To get a first idea of n_k, we use the variance ratio criterion (VRC) of Caliński and Harabasz (1974), which considers the relationship between the within-group average squared distances and the between-group average squared distances. Additionally, we inspect other reasonable values for n_k, for example, considering $n_k \pm 2$ after n_k is established using the VRC. This flexibility choosing n_k is helpful, for instance, to include a priori information or empirical knowledge in the interpretation. As the result of cluster analysis, we obtain n_k clustered families Fi (where $i = 1, 2, \ldots, n_k$) of models and corresponding sets of residuals denoted by $\boldsymbol{M}_{Fi}$ and $\boldsymbol{\delta}_{Fi}$, respectively.

2.5. Analyzing sets of models

Different statistical measures can summarize the information from a set of models (see the last two steps in Fig. 1). For example, we derive a representative (i.e., central tendency) model from a set of models $\boldsymbol{M}$ (e.g., $\boldsymbol{M}_{acc}$, $\boldsymbol{M}_{F1}$, or $\boldsymbol{M}_{F2}$) by calculating the median model $\mu_{1/2}(\boldsymbol{M})$ and estimate the variability within the set by the interquartile range $IQR(\boldsymbol{M})$ as used, for example, by Tronicke et al. (2012) and Fernández-Martínez et al. (2012). We want to highlight that when using robust statistics (e.g.,

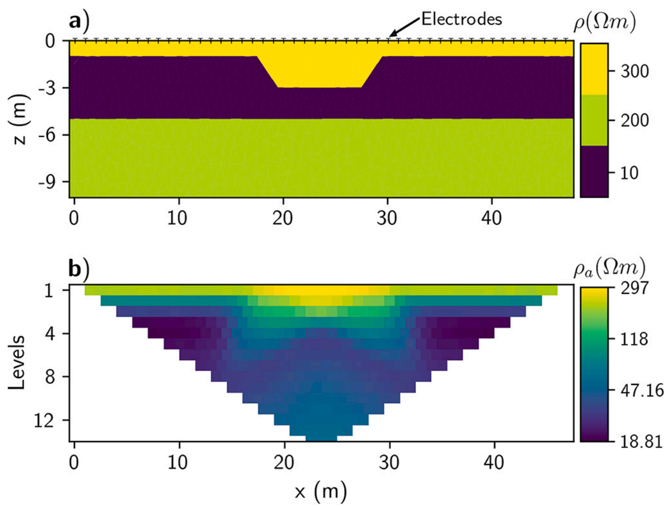

Fig. 2. (a) Synthetic input resistivity model indicating electrode positions along the surface and (b) synthetic apparent resistivity ρ_a data visualized as a Wenner pseudo-section.

median), we are less affected by outliers than when using non-robust statistics (e.g., mean) and, thus, less sensitive to the thresholds defined in the model selection process (Section 2.2). Additionally, uncertainties for a specific resistivity range are assessed using probability maps as suggested, for example, by Hermans et al. (2015), Fernández-Martínez et al. (2017), and De Pasquale et al. (2019). Here, we estimate the probability P of a certain range of resistivities $P(\boldsymbol{M}, [\rho_{min}, \rho_{max}])$ by counting the cells in $\boldsymbol{M}$ falling in the specified resistivity range and dividing it by the number of models in $\boldsymbol{M}$. This interpretation tool is useful if there is a well-defined geophysical target; i.e., a layer in the subsurface characterized by well-constrained resistivity values. These statistical analyses are performed in a pixel-wise fashion considering that all models have the same discretization (Section 2.3).

A critical point when working with resistivity models and values, respectively, is to define whether the statistical calculations (e.g., $\mu_{1/2}(\boldsymbol{M})$ and $IQR(\boldsymbol{M})$) are performed on a linear or logarithmic scale. We perform a preliminary statistical analysis using both scales (also for our later shown examples) to decide which scale is more appropriate. We find that calculations using a linear scale typically favor smooth solutions. In contrast, we notice that when using a logarithmic scale, we favor more blocky solutions. Because in our examples we aim to image subsurface interfaces using a layer-based parameterization, we decide to use the logarithmic scale to perform all calculations needed by our statistical analyses.

2.6. Analyzing sets of data residuals

A careful analysis of the data residuals (see Eq. (4)) can provide further insights into the solutions of the formulated inverse problem; for example, to identify any systematic effects, outliers, or biases that may distort our inversion results (e.g., Zhou and Dahlin, 2003; Juhojuntti and Kamm, 2015; Constable et al., 2015). In this study, we derive the central tendency and dispersion from a set of data residuals $\boldsymbol{\delta}$ (e.g., $\boldsymbol{\delta}_{acc}$, $\boldsymbol{\delta}_{F1}$, or $\boldsymbol{\delta}_{F2}$) by a pixel-wise calculation of the median of the residuals $\mu_{1/2}(\boldsymbol{\delta})$ and the interquartile range of the residuals $IQR(\boldsymbol{\delta})$, respectively. Additionally, in the same way as discussed in Section 2.2, we globally evaluate the statistical descriptors $RMSLE(\boldsymbol{\delta})$, $IQR(\boldsymbol{\delta})$, and $q_{90}(\boldsymbol{\delta})$ because they are useful to make comparisons between the clustered residuals (see the last two steps in Fig. 1). Finally, we also want to highlight that the clustered families may be further filtered using our model selection strategy (Section 2.2).

3. Case studies

We use two 2D ERT data sets to illustrate our workflow (Fig. 1). The first one is a synthetic example where we know the underlying input resistivity model and the noise level of the data. The second one is a field

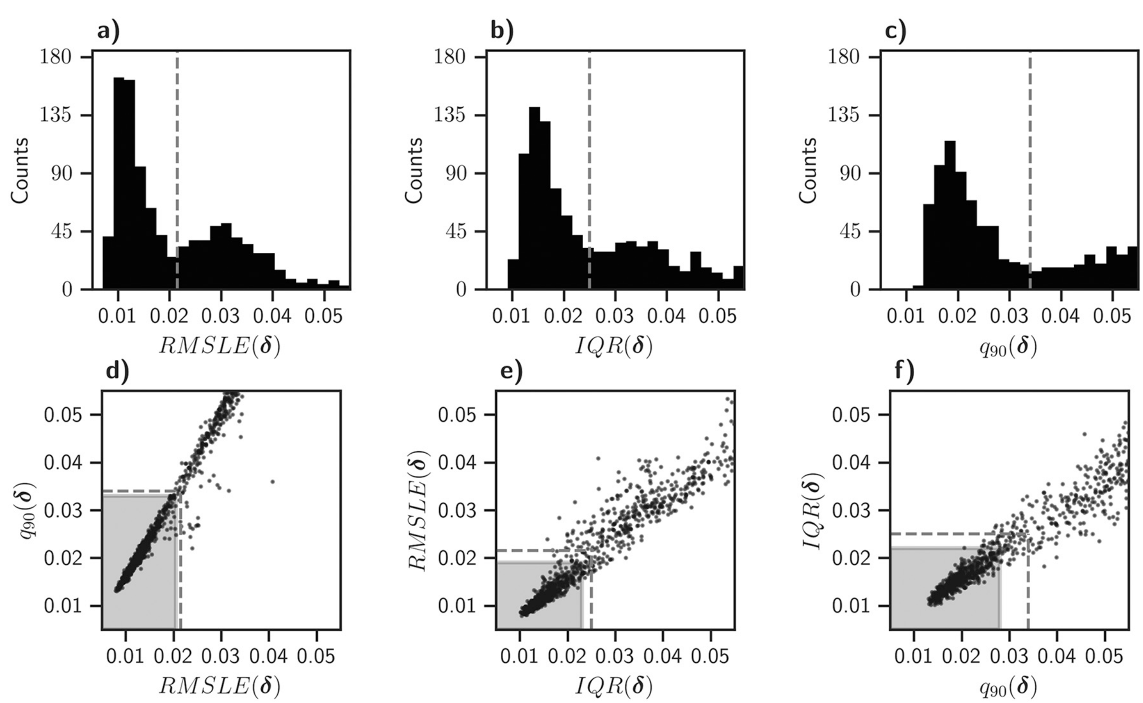

Fig. 3. Determination of the acceptance thresholds using $RMSLE(\boldsymbol{\delta})$, $IQR(\boldsymbol{\delta})$, and $q_{90}(\boldsymbol{\delta})$ to filter the initial model ensemble of our synthetic example. (a-c) Histograms indicating by the grey dashed lines the first estimation of the acceptance thresholds. (d-f) Scatter plots indicating by the grey rectangles the refined and final threshold values (note that the grey dashed lines highlight the thresholds defined by analyzing the histograms). For visualization purposes, we limit all axis to fits up to 0.055.

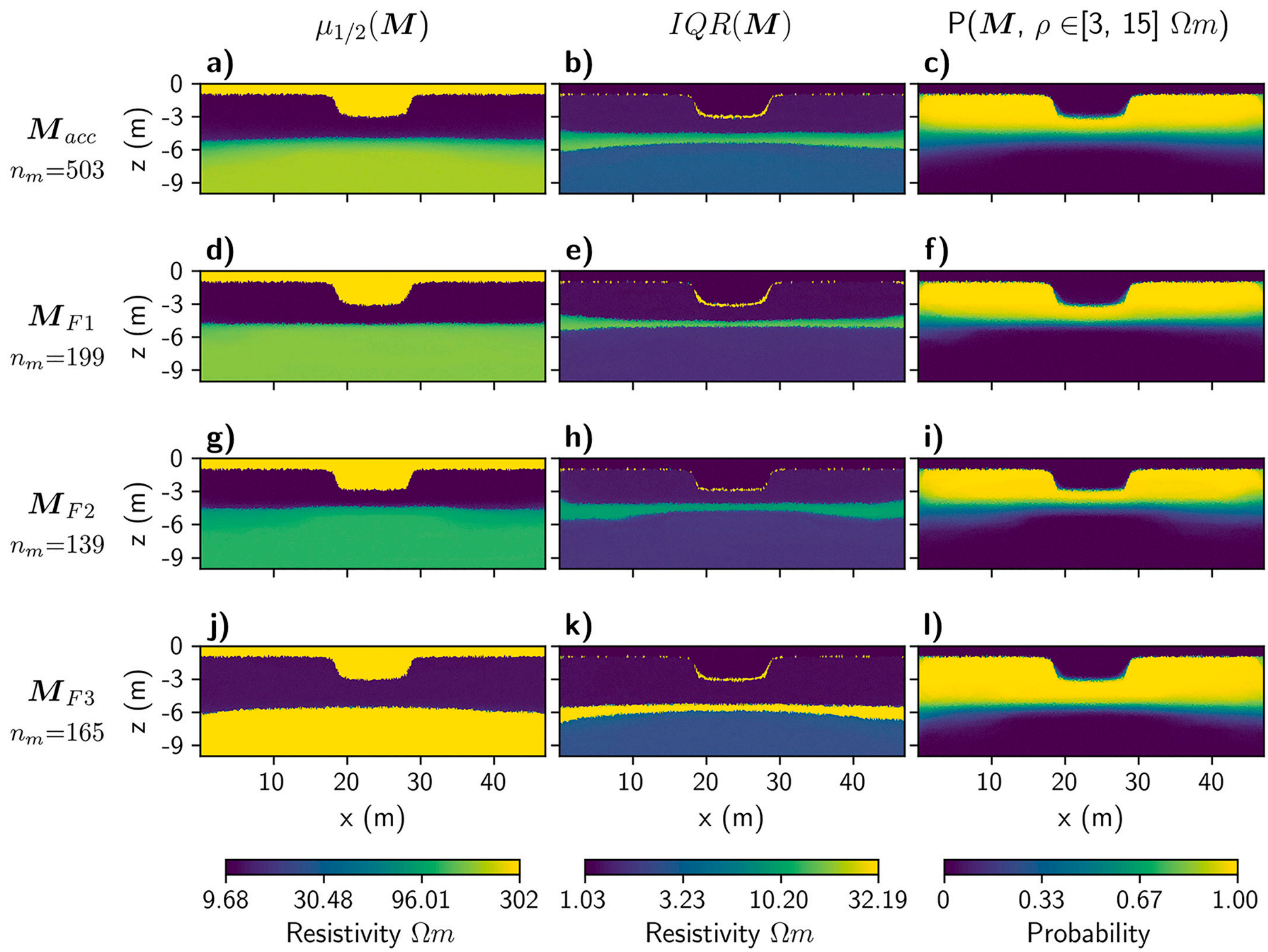

Fig. 4. Representation of the pixel-wise calculation of the $\mu_{1/2}(\boldsymbol{M})$, $IQR(\boldsymbol{M})$, and $P(\rho \in [3, 15]\Omega m)$ for our synthetic example. (a-c) statistics for the accepted models $\boldsymbol{M}_{acc}$ and (d-l) for the clustered families $\boldsymbol{M}_{F1}$, $\boldsymbol{M}_{F2}$, and $\boldsymbol{M}_{F3}$. n_m represents the size of each ensemble of models.

example where our geological background knowledge suggests a layered subsurface consisting of sedimentary layers with different pore fillings.

3.1. Synthetic example

Our synthetic input model consists of three layers (Fig. 2a). The uppermost layer represents a channel fill-like structure with $\rho = 300\ \Omega m$, the middle layer is a conductive layer with $\rho = 10\ \Omega m$, and the lowermost layer is characterized by $\rho = 200\ \Omega m$. To obtain the resistivity response of this model, we employ the BERT forward solver (Günther et al., 2006; Rücker et al., 2006) using a Wenner array configuration and 48 electrodes equally spaced in 1 m intervals along the surface. The simulated data set is contaminated by adding 0.01 of Gaussian noise (in terms of our objective function Eq. (3)) to simulate small-scale variabilities and noise as often observed in field data sets. In total, we obtain a pseudo-section with 360 apparent resistivity measurements as shown in Fig. 2b.

We invert our synthetic data using the layer-based global inversion strategy as described in Section 2.1. We use three interfaces (four layers) with four nodes for each interface resulting in a total of 43 model parameters (see Eq. (2)). To perform the global inversion, we use the PSO-based inversion strategy with a swarm size of 88 particles and save the best model after 400 iterations. These PSO parameters, as well as the model parameterization, are selected after some initial parameter testing. We repeated the inversion until we got a representative ensemble consisting of 1103 models.

To evaluate and filter the generated ensemble (Section 2.2), we explore possible threshold values using histograms and scatter plots of $RMSLE(\boldsymbol{\delta})$, $IQR(\boldsymbol{\delta})$, and $q_{90}(\boldsymbol{\delta})$ as shown in Fig. 3. From the histograms in Fig. 3a-c we notice a bimodal distribution in all variables. In order to reject models where the optimization may not have converged, we define upper thresholds (see the dashed vertical lines in Fig. 3a-c) to exclude models with rather high $RMSLE(\boldsymbol{\delta})$, $IQR(\boldsymbol{\delta})$, and $q_{90}(\boldsymbol{\delta})$ values, respectively. Then, we refine these threshold values using the scatter plots shown in Fig. 3d-f to select the models belonging to the main cloud of points. In Fig. 3d-f, we consider the smallest value of the statistical descriptor to define our model selection criteria. In the end, we accept models satisfying $RMSLE(\boldsymbol{\delta}) \leq 0.019$ (Fig. 3e), $IQR(\boldsymbol{\delta}) \leq 0.022$ (Fig. 3f), and $q_{90}(\boldsymbol{\delta}) \leq 0.028$ (Fig. 3f) resulting in 503 accepted models and residual vectors representing the ensembles $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$, respectively. Furthermore, because the remaining 503 models in $\boldsymbol{M}_{acc}$ have different discretizations containing 4600 ± 750 cells, we interpolate them to a finer common mesh with 16,502 cells (see Section 2.3).

Having all models with the same discretization, we calculate $\mu_{1/2}(\boldsymbol{M}_{acc})$, $IQR(\boldsymbol{M}_{acc})$, and $P(\boldsymbol{M}_{acc}, \rho \in [3, 15]\Omega m)$ to get a first impression regarding the characteristics of $\mathbf{M}_{acc}$ (Fig. 4a-c). Although the median model (Fig. 4a) represents a reasonable solution compared to the input model (Fig. 2a), we notice (when analyzing individual models) that some models show significant deviations from this median model; especially for depths $z < 5$ m. Thus, to further analyze $\boldsymbol{M}_{acc}$ and to

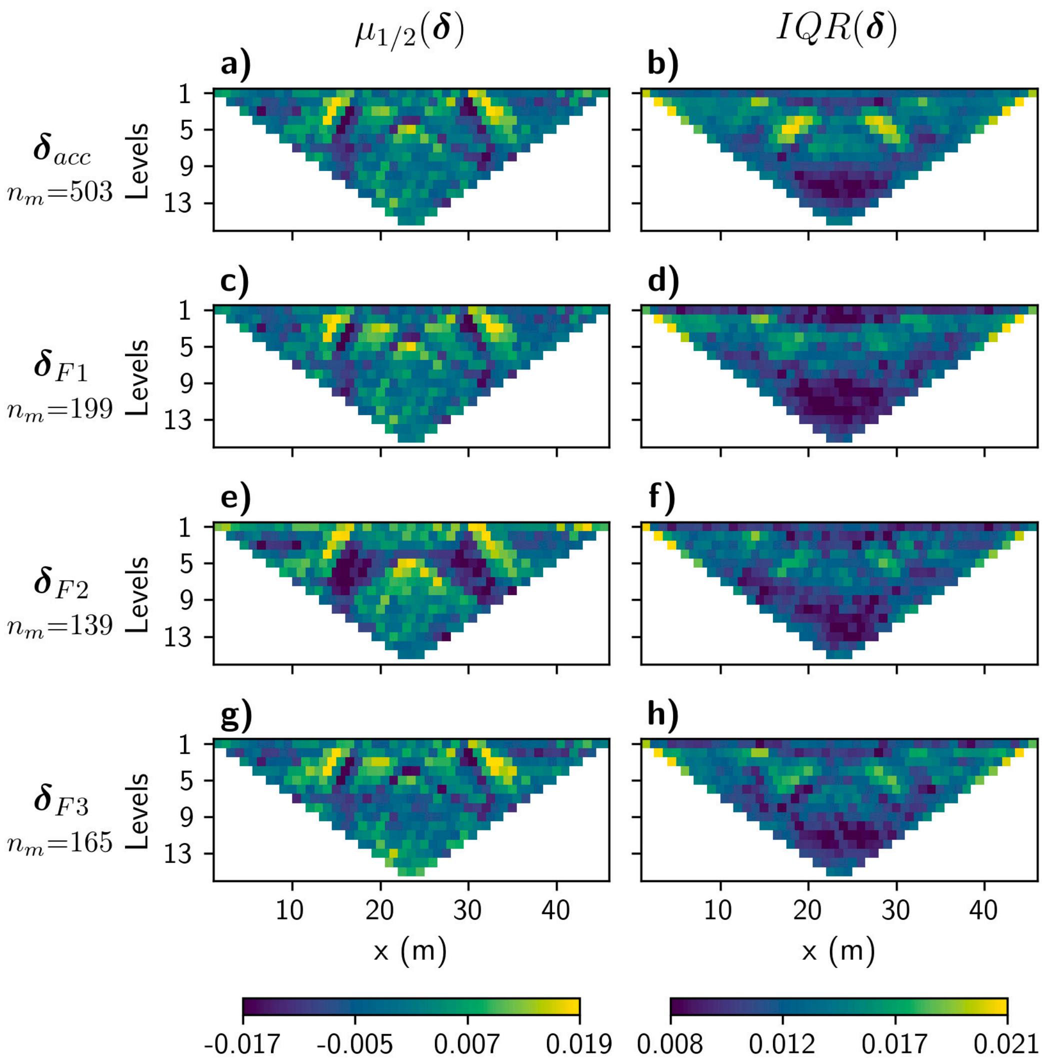

Fig. 5. Pseudo-section representation of the pixel-wise calculation of $\mu_{1/2}(\boldsymbol{\delta})$ and $IQR(\boldsymbol{\delta})$ for our synthetic example. (a-b) statistics for the accepted data residuals $\boldsymbol{\delta}_{acc}$ and (c-h) for the clustered families $\boldsymbol{\delta}_{F1}$, $\boldsymbol{\delta}_{F2}$, $\boldsymbol{\delta}_{F3}$. n_m represents the size of each ensemble of residual vectors.

identify potential groups of similar models, we perform cluster analysis taking into account $\boldsymbol{M}_{acc}$ and $\boldsymbol{\delta}_{acc}$ (see Section 2.4). Because each model in $\boldsymbol{M}_{acc}$ significantly differs in size to the corresponding residuals in $\boldsymbol{\delta}_{acc}$ (16,502 versus 360 values, respectively), we resample each model in $\boldsymbol{M}_{acc}$ by extracting 18 vertical resistivity profiles (up to a depth of $z = 10$ m) located between $x = 8$ m and $x = 40$ m with a lateral spacing of 1.88 m resulting in 771 values for each model. Moreover, we standardize the resampled model values and residuals and define $\boldsymbol{K}$ according to Eq. (5). Using the VRC and visual inspection, we found that $n_k = 3$ (i.e., three model families F1, F2, and F3) is sufficient to describe the variability within our reduced model ensemble. With this, 199 models are assigned to F1, 139 to F2, and 165 to F3.

The central tendency, dispersion, and probability models for the clustered model families *Fi* are shown in Fig. 4d-l. All median models $\mu_{1/2}(\boldsymbol{M}_{Fi})$ (Fig. 4d, g, and j) indicate a three-layer solution. However, when comparing these models in more detail, we notice some differences in the structures of the lower boundary of the conductive layer and in the resistivity of the lowermost layer. Also the $IQR(\boldsymbol{M}_{Fi})$ models (Fig. 4e, h, and k) indicate a three-layer case and highlight the variabilities around the layer interfaces. Because the conductive layer is the one controlling the resulting model structure (also our hypothetical target), we further investigate the uncertainties associated with this layer using the probability map $P(\boldsymbol{M}_{Fi}, \rho \in [3,15]\Omega m)$ (Fig. 4f, i, and l). These maps show high probabilities on both sides of and below the central channel-like structure, and the probability decreases as we approach the bottom of the conductive layer.

Furthermore, we analyze the corresponding sets of data residuals in a pixel-wise and global fashion (as outlined in Section 2.6). The pixel-wise representations of the central tendency and dispersion for all residuals $\boldsymbol{\delta}_{acc}$ and for each family of residuals $\boldsymbol{\delta}_{Fi}$ are shown in Fig. 5. The $\mu_{1/2}(\boldsymbol{\delta}_{Fi})$ pseudo-sections present similar patterns as $\mu_{1/2}(\boldsymbol{\delta}_{acc})$. This indicates that the models in these families fitted individual data points in a similar fashion, where the highest residuals are observed along diagonal lines around 18 and 30 m in the first levels of the pseudo-sections. On the other hand, all $IQR(\boldsymbol{\delta}_{Fi})$ pseudo-sections show similar variability patterns, while the $IQR(\boldsymbol{\delta}_{acc})$ shows slightly higher values, especially in the central part. In Fig. 6, we additionally illustrate the global evaluation of the *RMSLE*, *IQR*, and q_{90} for $\boldsymbol{\delta}_{acc}$ and for all $\boldsymbol{\delta}_{Fi}$ by histograms. We see that the histograms of $\boldsymbol{\delta}_{F1}$ and $\boldsymbol{\delta}_{F3}$ show a similar shape compared to the histogram of $\boldsymbol{\delta}_{acc}$ in all three statistical measures. However, for $\boldsymbol{\delta}_{F2}$, the histograms of these measures are different and slightly shifted towards higher values (compared to all other histograms). Overall, this residual analysis indicates that the three families fit the data to equivalent levels; i.e., if this was an actual field example, we could not distinguish the best family without any further information. However, if we consider the information of the known input model (Fig. 2a), we can conclude that F1 (Fig. 4d) represents the best family of models to solve our inverse problem. Although in this example we knew the noise level of our data,

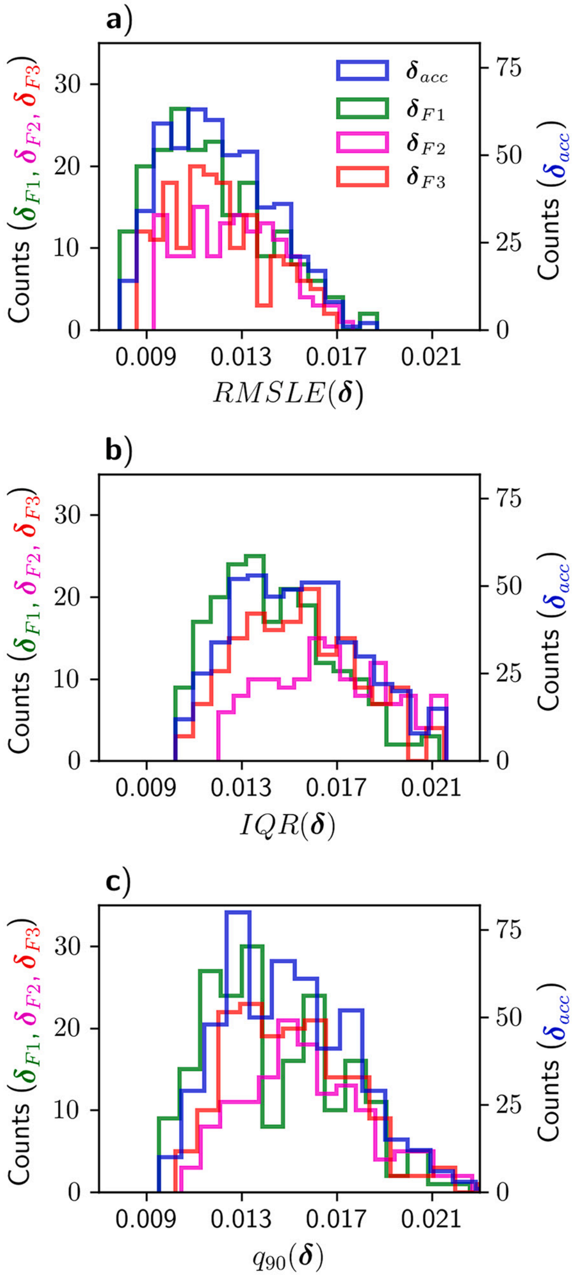

Fig. 6. (a-c) Histograms of the global statistical descriptors $RMSLE(\delta)$, $IQR(\delta)$, and $q_{90}(\delta)$ for the synthetic example. Notice that the corresponding counts of δ_{F1}, δ_{F2}, and δ_{F3} are represented by the left axis, and of δ_{acc} by the right axis. The legend in (a) also applies for (b) and (c).

we decide to work with models that fit the data at lower and higher values than the added noise level. This avoids focusing on models suspected to under-fitting or over-fitting the data and, thus, ensures more detailed insights of the inverse problem (e.g., Vasco et al., 1993).

3.2. Field example

In September 2019, we performed a 2D ERT survey in a coastal environment west of the town Loon-Plage, Northern France. This experiment was performed during an archaeological site evaluation by the French national institute for preventive archaeological research (INRAP), in the framework of the Port of Dunkirk extension project. To acquire our ERT data, we used 96 electrodes with a regular spacing of 2 m and collected our data using the Wenner array configuration. A single roll-along was performed by moving 50% of the electrodes. In the end, we recorded 2501 apparent resistivities along a 284 m long profile. After a first data quality evaluation, we removed the first level (186 data points) from the pseudo-section to reduce the influence of near-surface soil heterogeneities because the goal of this survey was to characterize deeper geological structures as needed to develop a hydrogeological understanding of this field site. In Fig. 7, we present the measured raw data.

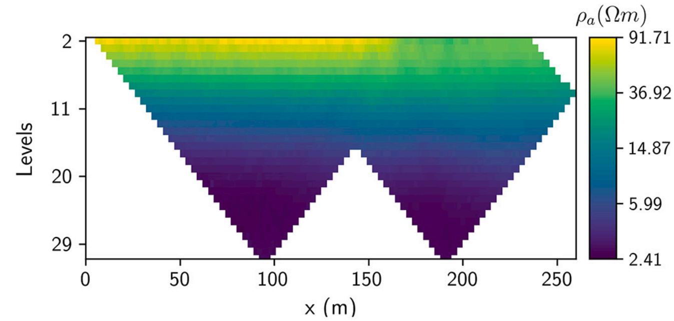

Fig. 7. Raw field data recorded at our field site in Loon-Plage, France using the Wenner array configuration.

Following our workflow, we invert our field data set using our layer-based model parameterization which is consistent with preliminary geological information of the study area. For example, Houthuys et al. (1993) describe the local geology as a transitional environment, where the first 15 to 25 m of the subsurface is characterized by horizontally stratified sandy Holocene sediments with clayey and peaty intercalations. Considering this a priori information, we invert our data set using four interfaces and three nodes for each interface, resulting in 45 model parameters (see Eq. (2)). In the PSO, we use 50 particles and a stopping criterion of a maximum of 400 iterations. In the end, we generate an ensemble of 802 models using this inversion strategy.

Similarly to our synthetic study, we use the histograms and scatter plots shown in Fig. 8 to obtain the thresholds values $t1$, $t2$, and $t3$ to remove models from the ensemble representing non-converged inversions runs (Section 2.2). As indicated in the plots of Fig. 8, we accept models satisfying $RMSLE(\delta) \leq 0.027$, $IQR(\delta) \leq 0.03$, and $q_{90}(\delta) \leq 0.038$ resulting in a reduced ensemble consisting of 744 models and residuals ($\boldsymbol{M}_{acc}$ and δ_{acc}, respectively). Furthermore, because the remaining 744 models in $\boldsymbol{M}_{acc}$ have different discretizations containing 4700 $\pm$ 700 cells, we interpolate them to a finer common mesh with 75,679 cells (see Section 2.3).

Having all models with the same discretization, we calculate $\mu_{1/2}(\boldsymbol{M}_{acc})$, $IQR(\boldsymbol{M}_{acc})$, and $P(\boldsymbol{M}_{acc}, \rho \in [1,3]\Omega m)$ as shown in Fig. 9a-c. From $\mu_{1/2}(\boldsymbol{M}_{acc})$ model (Fig. 9a), we learn that four layers (i.e., three interfaces) are sufficient to characterize the subsurface; i.e., two uppermost layers of approximately five meters thickness with a lateral contact at $x \approx 160$ m, a layer at depths between $z \approx 5$ m and $z \approx 10$ m, and a lowermost layer below $z \approx 10$ m. Furthermore, from the $IQR(\boldsymbol{M}_{acc})$ model (Fig. 9b), we notice that the dominant variability is found at $z \approx 10$ m, but there is also some variability associated to the overlying layer at depths between $z \approx 5$ m and $z \approx 10$ m. Additionally, when analyzing individual models (not shown here), we notice several models showing a layer at depths below 20 m with slightly increased resistivities compared to the overlying layer. To study this structure in more detail, we calculate $P(\boldsymbol{M}_{acc}, \rho \in [1,3]\Omega m)$ as shown in Fig. 9c where this resistivity range is defined after inspecting in more detail typical resistivities values (e.g., using histograms) of the deeper layers ($z < 10$ m) for each ensemble. Note that choosing different resistivity ranges may be motivated by a well-defined geophysical target, which in our case is a saltwater-saturated layer characterized by the lowest resistivity values in our inverted models. Alternatively, we could have used other resistivity ranges, for example, $P(\boldsymbol{M}_{Fi}, \rho > 40 \quad \Omega m)$ to study the near-

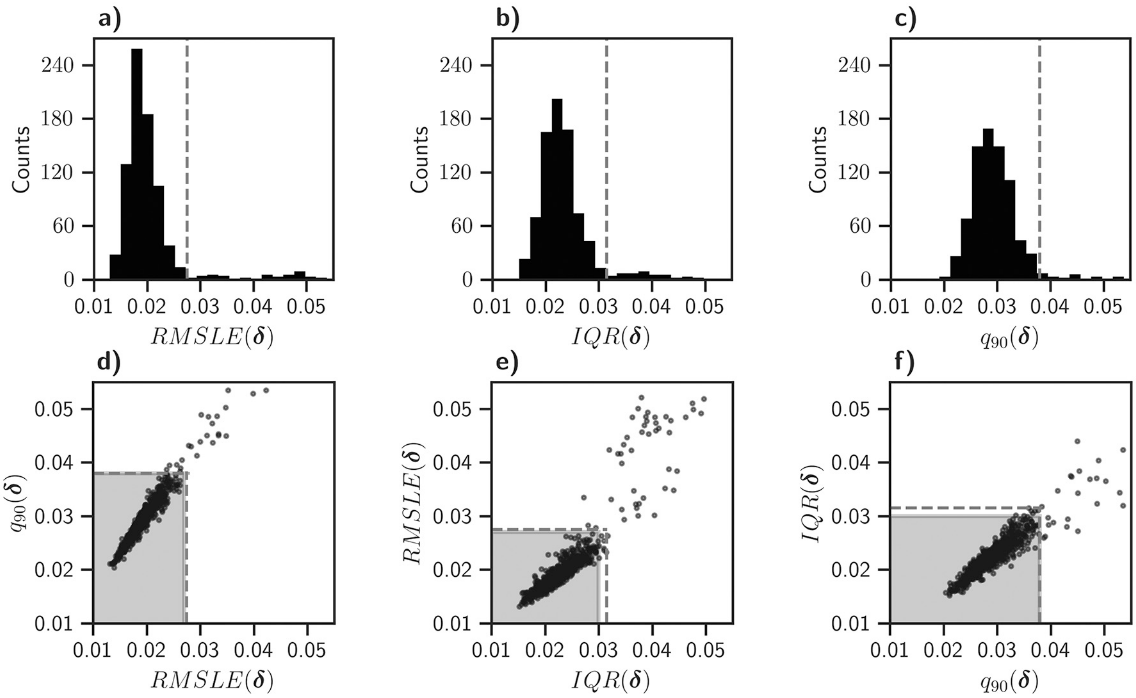

Fig. 8. Determination of the acceptance thresholds using $RMSLE(\delta)$, $IQR(\delta)$, and $q_{90}(\delta)$ to filter the initial model ensemble of our field example. (a-c) Histograms indicating by the grey dashed lines the first estimation of the acceptance thresholds. (d-f) Scatter plots indicating by the grey rectangles the refined and final threshold values (note that the grey dashed lines highlight the thresholds defined by analyzing the histograms). For visualization purposes, we limit all axis to fits up to 0.055.

surface structures in more detail if these model features are of interpretational interest. From the probability map, we notice, at depths between $z \approx 10$ m and $z \approx 23$ m, a probability of $\sim$1. However, this probability slightly decreases below $z \approx 23$ m. The analysis of individual models and Fig. 9c indicates the existence of different groups of models describing different geological scenarios difficult to be identified by analyzing Fig. 9a-c. To further explore this, we perform cluster analysis as outlined in Section 2.4.

Before clustering, the models $\boldsymbol{M}_{acc}$ are resampled along 50 vertical resistivity profiles up to a depth of $z = 30$ m, laterally distributed between $x = 20$ m and $x = 260$ m with an equal spacing of 4.9 m. In total, we extract 3556 values for each model, which is comparable to the number of data and residuals points, respectively. Then, we standardize the resampled models and residual values to define $\boldsymbol{K}$ according to Eq. (5). In cluster analysis, by using the VRC and visual inspections, we find that three families (F1, F2, and F3) are sufficient to explain the data and to obtain a reasonable solution. Finally, we thus obtain a solution with 258 models in F1, 119 in F2, and 367 in F3.

The statistics of the clustered model families Fi are shown in Fig. 9d-l. The median models $\mu_{1/2}(\boldsymbol{M}_{Fi})$ (Fig. 9d, g, and j) suggest three different subsurface scenarios. For example, the median model $\mu_{1/2}(\boldsymbol{M}_{F1})$ (Fig. 9d) represents a resistive layer above $z \approx 10$ m including a $\sim$5 m thick more conductive top layer for $x > 160$ m. In contrast, $\mu_{1/2}(\boldsymbol{M}_{F2})$ (Fig. 9g) and $\mu_{1/2}(\boldsymbol{M}_{F3})$ (Fig. 9j) represent a similar near-surface-structure, where above $z \approx 10$ m there is a more conductive layer embedding a more resistive top layer for $x < 160$ m. However, $\boldsymbol{M}_{F2}$ also shows a resistivity contrast at $z \approx 23$ m. Additionally, we infer the uncertainties associated with the layer interfaces in $\mu_{1/2}(\boldsymbol{M}_{Fi})$ calculating the $IQR(\boldsymbol{M}_{Fi})$ models (Fig. 9e, h, and k). From the IQR models, we learn that the main uncertainties are associated with the layer at $z \approx 10$ m, where, for example, we notice higher variations for $\boldsymbol{M}_{F1}$ than for $\boldsymbol{M}_{F2}$ and $\boldsymbol{M}_{F3}$. Furthermore, $\boldsymbol{M}_{F2}$ contains the highest increased variability for the layer at $z \approx 23$ m. Finally, we use $P(\boldsymbol{M}_{Fi}, \rho \in [1,3]\Omega m)$ to better explore the models at depths below $z \approx 10$ m (Fig. 9f, i, and l). Here, we observe that $\boldsymbol{M}_{F2}$ is most different compared with $\boldsymbol{M}_{F1}$ and $\boldsymbol{M}_{F3}$. This is in agreement with Fig. 9g, where we notice that the resistivity increases at depths below $z \approx 23$ m. In the hypothetical case that we are asked to identify the resistivity range that maximize the probability of the conductive layer in $\boldsymbol{M}_{F2}$, we would have to modify iteratively the applied resistivity range (e.g., trying $P(\boldsymbol{M}_{F2}, \rho \in [3,8]\Omega m)$, $P(\boldsymbol{M}_{F2}, \rho \in [5,10]\Omega m)$, ...). Overall, we notice significant differences between the three families, which is illustrating the benefit of performing such a cluster analysis step.

To complement our ensemble analysis, we calculate and examine the corresponding sets of data residuals (for all residuals $\boldsymbol{\delta}_{acc}$ and each family of residuals $\boldsymbol{\delta}_{Fi}$) as outlined in Section 2.6. In the pixel-wise analysis of the residuals (Fig. 10), we observe that the $\mu_{1/2}(\boldsymbol{\delta}_{F1})$ and $\mu_{1/2}(\boldsymbol{\delta}_{F3})$ pseudo-sections show similar overall patterns compared to the $\mu_{1/2}(\boldsymbol{\delta}_{acc})$ pseudo-section. However, $\mu_{1/2}(\boldsymbol{\delta}_{F2})$ is different and the corresponding pseudo-section (Fig. 10e) shows reduced residual values for $levels > 15$ compared to Fig. 10a, c, and g. Although the $\mu_{1/2}(\boldsymbol{\delta})$ pseudo-sections are indicating high relative error values ($0.04 - 0.05$) for the first and deeper levels, these values are still in an acceptable range for an ERT survey. Additionally, we find that the $IQR(\boldsymbol{\delta})$ pseudo-sections (Fig. 10b, d, f, and h) present the same main patterns for all sets of residuals ($\boldsymbol{\delta}_{acc}$, $\boldsymbol{\delta}_{F1}$, $\boldsymbol{\delta}_{F2}$, $\boldsymbol{\delta}_{F3}$) where the highest variabilities are located in the uppermost and lowermost levels. Furthermore, we notice a diagonal pattern originating in the uppermost level at $x \approx 160$ m. These diagonals are related to the location of the lateral resistivity contrast discussed above. Complementary to the pixel-wise analysis, we globally assess the $RMSLE$, IQR, and q_{90} for $\boldsymbol{\delta}_{acc}$ and $\boldsymbol{\delta}_{Fi}$ (Fig. 11). From these histograms, we learn that $\boldsymbol{\delta}_{F1}$ and $\boldsymbol{\delta}_{F3}$ are characterized by the highest values for all variables, while the histograms of $\boldsymbol{\delta}_{F2}$ are shifted towards lower values compared to $\boldsymbol{\delta}_{F1}$ and $\boldsymbol{\delta}_{F3}$. From the pixel-wise and global analysis of the residuals, we notice a better performance of F2, where it is important to highlight that F2 is the one indicating a layer boundary at a depth of $\sim$23 m (see Fig. 9g). This indicates that the data points collected at the lowest levels are still sensitive to the small resistivity contrasts of the models collected in F2.

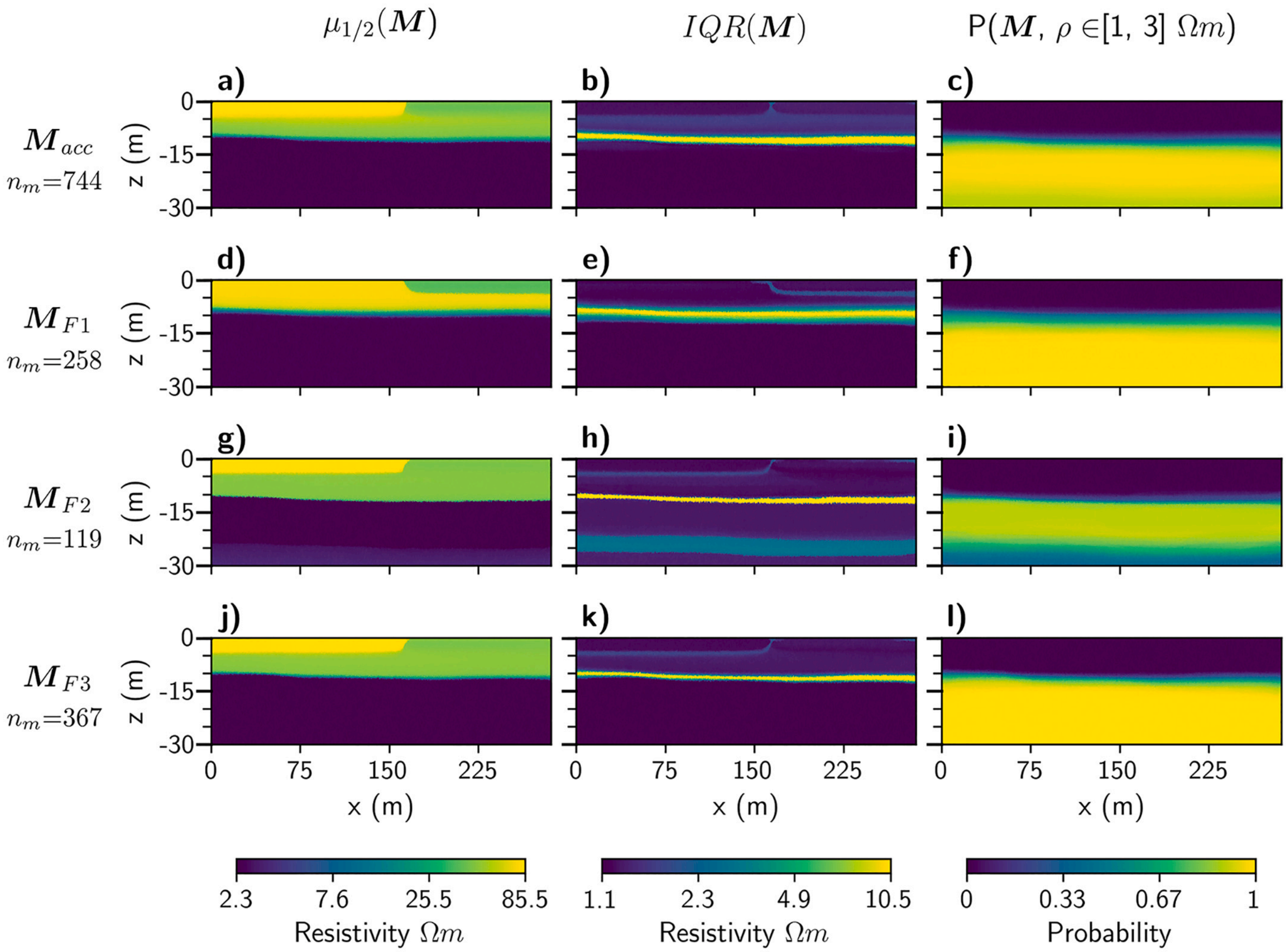

Fig. 9. Representation of the pixel-wise calculation of the $\mu_{1/2}(\boldsymbol{M})$, $IQR(\boldsymbol{M})$, and $P(\rho \in [3, 15]\Omega m)$ for our field example. (a-c) statistics for the accepted models $\boldsymbol{M}_{acc}$ and (d-l) for the clustered families $\boldsymbol{M}_{F1}$, $\boldsymbol{M}_{F2}$, and $\boldsymbol{M}_{F3}$. n_m represents the size of each ensemble of models.

4. Discussion

Inversion of ERT data is a non-linear and non-unique problem. However, using global optimization algorithms such as PSO, we can obtain an ensemble of equivalent solutions. This study compiles a workflow (Fig. 1) that allows us to extract the main structures needed by the data and to assess the variability within the ensembles including the underlying residual distributions. To demonstrate our approach we used two 2D ERT data sets, which have been inverted using PSO and a layer-based model parameterization.

Although we iteratively use a global-search approach to invert a 2D ERT data set, some optimization runs might not converge (e.g., Lomax and Snieder, 1995; Douma et al., 1996). Furthermore, defining a reliable stopping criterion is not straightforward, especially when there is no clear knowledge of the data noise level. When an approximation of the noise level is obtained by repeated measurements (e.g., Zhou and Dahlin, 2003), it is recommended to obtain models whose misfits values fall within a range contemplating smaller and larger misfit values and, thus, avoid the influence of data perturbation or modification of the inversion strategy (e.g., Vasco et al., 1993). We must consider that such an approach may result in considering models that are under-fitting or over-fitting the data. This is why in the second step of our workflow (Fig. 1), we propose a model selection criteria that not only considers the models exceeding a certain threshold in the objective function (*RMSLE* ($\boldsymbol{\delta}$) in this study) but also considers other statistical descriptors such as $q_{90}(\boldsymbol{\delta})$ as in Barboza et al. (2019) and $IQR(\boldsymbol{\delta})$. In our examples, thresholds were initially defined in such a way that models located in the high-value tails of the $RMSLE(\boldsymbol{\delta})$, $IQR(\boldsymbol{\delta})$, and $q_{90}(\boldsymbol{\delta})$ histograms are discarded. However, just using simple thresholding may result in inaccuracies associated, for example, with the bin size of the histograms. We found that scatter plots help to further refine the threshold levels (see Fig. 3 and Fig. 8). Overall, the proposed way of filtering and formulating the model selection criteria allows for a detailed and robust selection of models from a generated ensemble.

To study model ensembles, some examples in the literature have successfully shown that central trends, dispersion measures, and probability maps are effective ways to summarize the information contained in an ensemble including the communication of uncertainties (e.g., Tronicke et al., 2012; Fernández-Martínez et al., 2012; De Pasquale et al., 2019). Such global ensemble statistics are useful if similar models compose the ensemble. This can be inspected, for example, by visually comparing the calculated median model with individual models from the ensemble. If the ensemble is rather heterogeneous, the median model might not be representative for most models and, thus, cluster analysis can be considered as a useful tool to analyze the ensemble in more detail. In this study, we extend previous ideas by using a cluster strategy that not only considers the model values (e.g., Vasco et al., 1993, 1996; Ramirez et al., 2005; Fernández-Martínez et al., 2017) but also the corresponding data residuals. This strategy results in clusters of similar models and ensures that they fit the data in a similar fashion (as

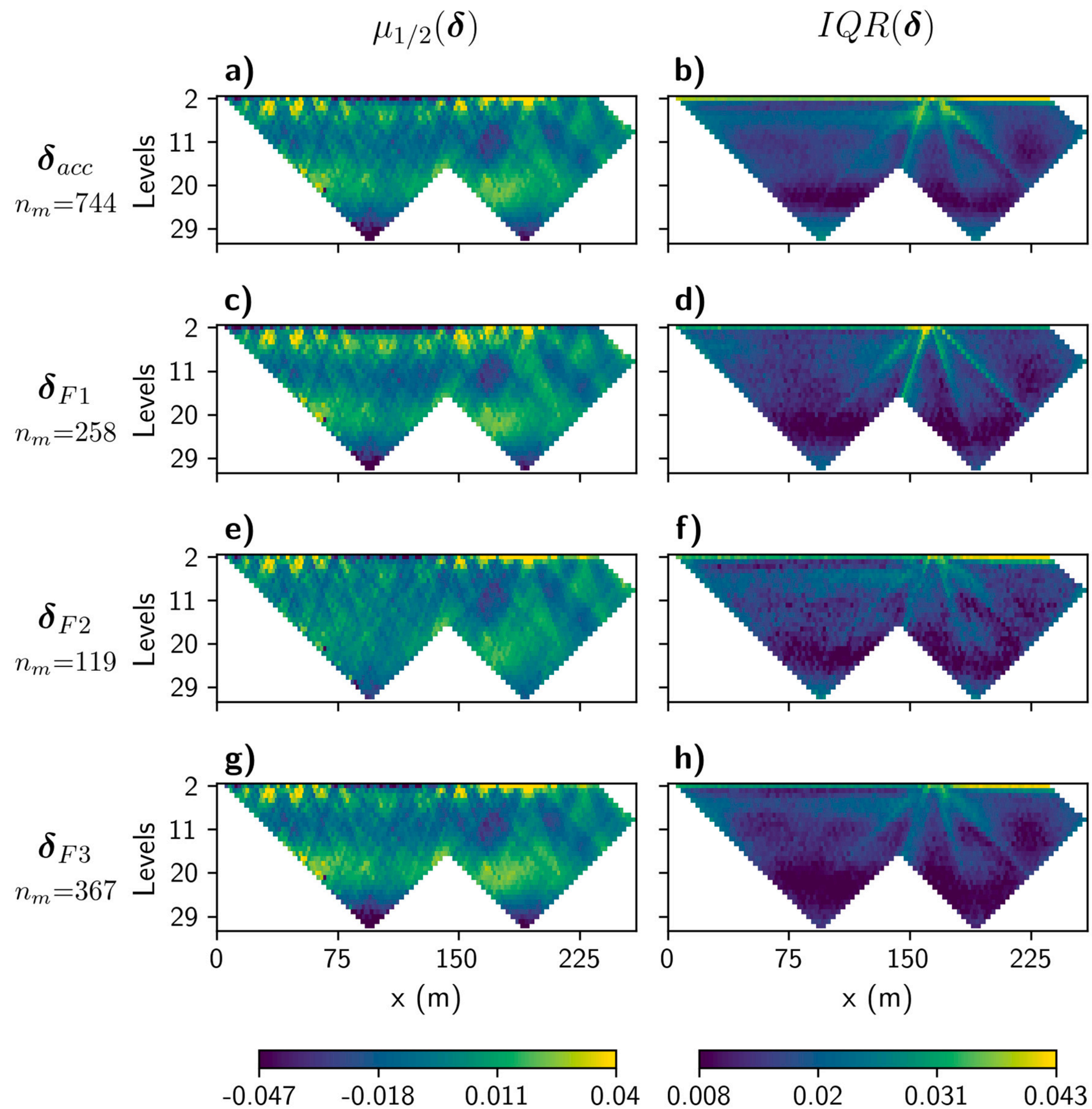

Fig. 10. Pseudo-section representation of the pixel-wise calculation of $\mu_{1/2}(\boldsymbol{\delta})$ and $IQR(\boldsymbol{\delta})$ for our field example. (a-b) statistics for the accepted data residuals $\boldsymbol{\delta}_{acc}$ and (c-h) for the clustered families $\boldsymbol{\delta}_{F1}$, $\boldsymbol{\delta}_{F2}$, $\boldsymbol{\delta}_{F3}$. n_m represents the size of each ensemble of residual vectors.

shown by the two examples in this study). Further extensions of our strategy could be, for example, to apply a weighted *k*-means algorithm, where the sensitivity matrix, the configuration factors, and/or measurement errors are used to weight the model and residual values, respectively. Although in this study we focus on the classical *k*-means algorithm, other algorithms such as spectral or agglomerative clustering might be considered to perform the classification.

Apart from analyzing sets of models, studying the corresponding sets of data residuals adds further valuable information to better understand the clustered ensemble. Typically, global evaluations of the residuals are used to get an overall view of each ensemble's performance. In this study, we additionally use a pixel-wise representation of the residuals. This visualization strategy might be used, for example, to explore the areas where the fit performs better or worse, find outliers, or determine numerical errors. We also highlight that when using a layer-based model parameterization the spatial error distribution may create some characteristic patches which can be explained by the limitation of a layer-based approach to fit small-scale heterogenities. Using both global and pixel-wise evaluations of the residuals provides a more robust criterion to accept/reject a specific set of models and to evaluate or modify the used inversion strategy (e.g., to adapt the parameterization strategy or the objective function).

Understanding uncertainties associated with the non-uniqueness of the inverse problem is crucial to create, for example, probabilistic structural and petrophysical translations. Nevertheless, this kind of analysis requires an ensemble composed of up to hundreds of independent solutions. Generating such ensembles through global optimization algorithms is computationally demanding. However, with recent computer developments, the calculation time has been considerably reduced. Furthermore, different methods help to mine the parameter solution space and to increase the number of model solutions (e.g., Sambridge, 1999; Tompkins et al., 2011; Fernández-Martínez et al., 2017).

5. Conclusions

This paper presents a workflow to assess ensembles of 2D ERT inversion results generated using global optimization strategies such as genetic algorithms, different simulate annealing varieties, PSO (as implemented in this study), and others. We found, in a first step, that

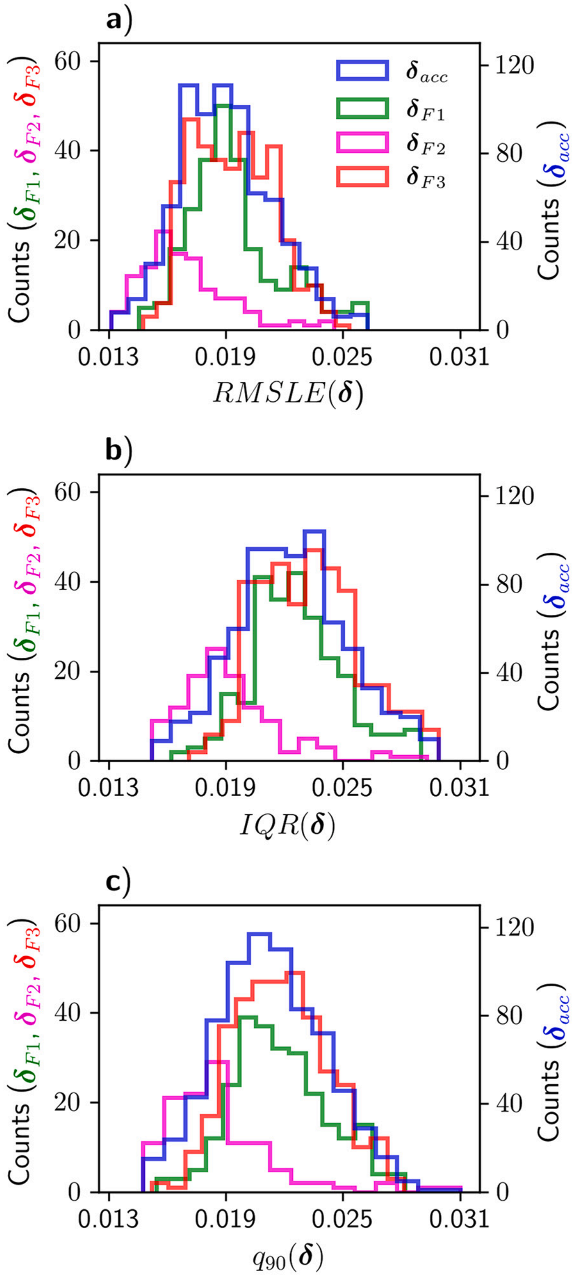

Fig. 11. (a-c) Histograms of the global statistical descriptors *RMSLE*(δ), *IQR*(δ), and $q_{90}(\delta)$ for the field example. Notice that the corresponding counts of $\boldsymbol{\delta}_{F1}$, $\boldsymbol{\delta}_{F2}$, and $\boldsymbol{\delta}_{F3}$ are represented by the left axis, and of $\boldsymbol{\delta}_{acc}$ by the right axis. The legend in (a) also applies for (b) and (c).

using the global evaluation of *RMSLE*(δ), *IQR*(δ), and $q_{90}(\delta)$ is a robust criterion to remove models that might not have completely converged and, thus, improve the posterior statistical characterization of the ensemble. Furthermore, the cluster analysis using the *k*-means algorithm considering both model and residual values shows to be an efficient and robust approach to better explore and understand our ensembles and, thus, allows us to identify different equivalent families of models with their corresponding data residuals. To summarize the information in such sets of clustered models, the central trend for each family can be treated as different plausible subsurface scenarios whose uncertainties can be derived by dispersion measurement statistics and probability maps. Additionally, analyzing the corresponding residual distributions for each family provides valuable information to further understand and assess the validity of the clustering output. In conclusion, our workflow eases and refines the interpretation of model ensembles resulting from the global inversion of 2D ERT data and might be easily adapted to other kinds of geophysical data and inverse problems, respectively.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank the German Academic Exchange Service (DAAD) for supporting Mauricio Arboleda Zapata [grant number 57395813]. This work was developed in the framework of the Research Training Group “Natural hazards and Risks in a Changing World” NatRiskChange 2043/2 (https://www.uni-potsdam.de/natriskchange/) funded by the Deutsche Forschungsgemeinschaft (DFG). We also thank the French national institute for preventive archaeological research (INRAP) for allowing us to perform the ERT survey in Loon-Plage and for logistical support during the fieldwork. Finally, we would like to thank the pyGIMLi crew for making the ERT processing routines available and providing online support.

References

Akça, I., Basokur, A.T., 2010. Extraction of structure-based geoelectric models by hybrid genetic algorithms. Geophysics 75 (1), F15–F22.
Auken, E., Christiansen, A.V., 2004. Layered and laterally constrained 2D inversion of resistivity data. Geophysics 69 (3), 752–761.
Barboza, F.M., Medeiros, W.E., Santana, J.M., 2018. Customizing constraint incorporation in direct current resistivity inverse problems: a comparison among three global optimization methods. Geophysics 83 (6), E409–E422.
Barboza, F.M., Medeiros, W.E., Santana, J.M., 2019. A user-driven feedback approach for 2D direct current resistivity inversion based on particle swarm optimization. Geophysics 84 (2), E105–E124.
Boiero, D., Socco, L.V., 2014. Joint inversion of Rayleigh-wave dispersion and P-wave refraction data for laterally varying layered models. Geophysics 79 (4). EN49–EN59.
Caliński, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Commun. Stat.-Theory Methods 3 (1), 1–27.
Constable, S., Orange, A., Key, K., 2015. And the geophysicist replied: “which model do you want?”. Geophysics 80 (3), E197–E212.
Cordua, K.S., Hansen, T.M., Mosegaard, K., 2012. Monte Carlo full-waveform inversion of crosshole GPR data using multiple-point geostatistical a priori information. Geophysics 77 (2), H19–H31.
de Groot-Hedlin, C.D., Vernon, F.L., 1998. An evolutionary programming method for estimating layered velocity structure. Bull. Seismol. Soc. Am. 88 (4), 1023–1035.
De Pasquale, G., Linde, N., Doetsch, J., Holbrook, W.S., 2019. Probabilistic inference of subsurface heterogeneity and interface geometry using geophysical data. Geophys. J. Int. 217 (2), 816–831.
Douma, H., Haney, M., 2013. Exploring nonlinearity and nonuniqueness in surface-wave inversion for near-surface velocity estimation. Lead. Edge 32 (6), 648–655.
Douma, H., Snieder, R., Lomax, A., 1996. Ensemble inference in terms of empirical orthogonal functions. Geophys. J. Int. 127 (2), 363–378.
Fernández-Martínez, J.L., García Gonzalo, E., Fernández Álvarez, J.P., Kuzma, H.A., Menéndez Pérez, C.O., 2010. PSO: a powerful algorithm to solve geophysical inverse problems application to a 1D-DC resistivity case. J. Appl. Geophys. 71 (1), 13–25.
Fernández-Martínez, J.L., Mukerji, T., García-Gonzalo, E., Suman, A., 2012. Reservoir characterization and inversion uncertainty via a family of particle swarm optimizers. Geophysics 77 (1), M1–M16.
Fernández-Martínez, J.L., Xu, S., Sirieix, C., Fernández-Muniz, Z., Riss, J., 2017. Uncertainty analysis and probabilistic segmentation of electrical resistivity images: the 2D inverse problem. Geophys. Prospect. 65, 112–130.
Furman, A., Ferre, T.P.A., Warrick, A.W., 2003. A sensitivity analysis of electrical resistivity tomography array Types using analytical element modeling. Vadose Zone J. 2 (3), 416–423.
Günther, T., Rücker, C., Spitzer, K., 2006. Three-dimensional modelling and inversion of DC resistivity data incorporating topography - II. Inversion. Geophys. J. Int. 166 (2), 506–517.
Hermans, T., Nguyen, F.F.F., Caers, J., 2015. Uncertainty in training image-based inversion of hydraulic head data constrained to ERT data: Workflow and case study. Water Resour. Res. 51 (7), 5332–5352.
Houthuys, R., De Moor, G., Somme, J., 1993. Shaping of the French-Belgian North Sea Coast throughout recent geology and history. In: Coastal Zone: Proceedings of the Symposium on Coastal and Ocean Management, pp. 27–40.
Juhojuntti, N., Kamm, J., 2015. Joint inversion of seismic refraction and resistivity data using layered models-applications to groundwater investigation. Geophysics 80(1): EN43–EN55.

Loke, M.H., Chambers, J.E., Rucker, D.F., Kuras, O., Wilkinson, P.B., 2013. Recent developments in the direct-current geoelectrical imaging method. J. Appl. Geophys. 95 (2013), 135–156.
Lomax, A., Snieder, R., 1995. Identifying sets of acceptable solutions to non-linear, geophysical inverse problems which have complicated misfit functions. Nonlinear Process. Geophys. 3/4(2):222–227.
MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability. University of California Press, pp. 281–297.
Oldenburg, D.W., Li, Y., 1999. Estimating depth of investigation in DC resistivity and IP surveys. Geophysics 64 (2), 403–416.
Ramirez, A.L., Nitao, J.J., Hanley, W.G., Aines, R., Glaser, R.E., Sengupta, S.K., Dyer, K. M., Hickling, T.L., Daily, W.D., 2005. Stochastic inversion of electrical resistivity changes using a Markov Chain Monte Carlo approach. J. Geophys. Res. Solid Earth 110 (2), 1–18.
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.
Rothman, D.H., 1985. Nonlinear inversion, statistical mechanics, and residual statics estimation. Geophysics 50 (12), 2784–2796.
Roy, L., Sen, M.K., McIntosh, K., Stoffa, P.L., Nakamura, Y., 2005. Joint inversion of first arrival seismic travel-time and gravity data. J. Geophys. Eng. 2 (3), 277–289.
Rücker, C., Günther, T., Spitzer, K., 2006. Three-dimensional modelling and inversion of dc resistivity data incorporating topography - I. Modelling. Geophys. J. Int. 166 (2), 495–505.
Rücker, C., Günther, T., Wagner, F.M., 2017. pyGIMLi: an open-source library for modelling and inversion in geophysics. Comput. Geosci. 109 (2017), 106–123.
Rumpf, M., Tronicke, J., 2014. Predicting 2D geotechnical parameter fields in near-surface sedimentary environments. J. Appl. Geophys. 101 (2014), 95–107.
Rumpf, M., Tronicke, J., 2015. Assessing uncertainty in refraction seismic traveltime inversion using a global inversion strategy. Geophys. Prospect. 63 (5), 1188–1197.
Sambridge, M., 1999. Geophysical inversion with a neighbourhood algorithm–II. Appraising the ensemble. Geophys. J. Int. 138 (3), 727–746.
Scheidt, C.C., Li, L., Caers, J., 2018. Quantifying Uncertainty in Subsurface Systems. John Wiley and Sons.
Sen, M.K., Stoffa, P.L., 1991. Nonlinear one-dimensional seismic waveform inversion using simulated annealing. Geophysics 56 (10), 1624–1638.
Sen, M.K., Stoffa, P.L., 1996. Bayesian inference, Gibbs' sampler and uncertainty estimation in geophysical inversion. Geophys. Prospect. 44 (2), 313–350.
Sen, M.K., Stoffa, P.L., 2013. Global Optimization Methods in Geophysical Inversion, 2nd ed. Cambridge University Press.
Shaw, R., Srivastava, S., 2007. Particle swarm optimization: a new tool to invert geophysical data. Geophysics 72 (2), F75–F83.
Smith, T., Hoversten, M., Gasperikova, E., Morrison, F., 1999. Sharp boundary inversion of 2D magnetotelluric data. Geophys. Prospect. 47 (4), 469–486.
Stoffa, P.L., Sen, M.K., 1991. Nonlinear multiparameter optimization using genetic algorithms: inversion of plane-wave seismograms. Geophysics 56 (11), 1794–1810.
Tarantola, A., 2005. Inverse problem theory and methods for model parameter estimation. SIAM. https://doi.org/10.1137/1.9780898717921. Pages: ix + 339.
Tompkins, M.J., Fernández-Martínez, J.L., Alumbaugh, D.L., Mukerji, T., 2011. Scalable uncertainty estimation for nonlinear inverse problems using parameter reduction, constraint mapping, and geometric sampling: Marine controlled-source electromagnetic examples. Geophysics 76 (4), F263–F281.
Tronicke, J., Paasche, H., Böniger, U., 2012. Crosshole traveltime tomography using particle swarm optimization: a near-surface field example. Geophysics 77 (1), R19–R32.
Vasco, D.W., Johnson, L.R., Majer, E.L., 1993. Ensemble inference in geophysical inverse problems. Geophys. J. Int. 115 (3), 711–728.
Vasco, D.W., Peterson, J.E., Majer, E.L., 1996. Nonuniqueness in traveltime tomography: Ensemble inference and cluster analysis. Geophysics 61 (4), 1209–1226.
Zhou, B., Dahlin, T., 2003. Properties and effects of measurement errors on 2D resistivity imaging surveying. Near Surf. Geophys. 1 (3), 105–117.
Zhu, D., Gibson, R., 2018. Seismic inversion and uncertainty quantification using transdimensional Markov chain Monte Carlo method. Geophysics 83 (4), R321–R334.