Facies discrimination with ERT using a probabilistic methodology: effect of sensitivity and regularization

Hermans Thomas, Stanford University, School of Earth, Energy and Environment Sciences, Geological Department

Irving James, University of Lausanne, Institute of Earth Sciences, Applied and Environmental Geophysics Group

Corresponding author

Hermans Thomas, Stanford University, 450, Serra Mall, Building 320, Room 120, 94305 Stanford, California.

Previously at University of Liege, Applied Geophysics and FNRS, Brussels.

E-mail : thermans@stanford.edu; thomas.hermans@ulg.ac.be

Tel.: 1-650-723-2223

Preprint submitted to *Near Surface Geophysics*

Full version available at: https://dx.doi.org/10.3997/1873-0604.2016047

Abstract

Electrical resistivity tomography (ERT) has become a standard geophysical method in the field of hydrogeology, as it has the potential to provide important information regarding the spatial distribution of facies. However, inverted ERT images tend to be grossly smoothed versions of reality because of the regularization of the inverse problem. In this study, we use a probabilistic methodology based upon co-located measurements to assess the utility of ERT to identify hydrofacies in alluvial aquifers. With this methodology, ERT images are interpreted in terms of the probability of belonging to pre-defined hydrofacies. We first analyze through a synthetic study the ability of ERT to discriminate between different facies. As ERT data suffer from a loss of sensitivity with depth, we find that low sensitivity regions are more affected by misclassification. To counteract this effect, we adapt the probabilistic framework to include the spatially varying data sensitivity. We then apply our learning to a field case. For the latter, we consider two different regularization procedures. In contrast to the data sensitivity which affects the facies probability to a limited amount, the regularization can affect the probability maps more considerably because it has a strong influence on the spatial distribution of inverted resistivity. We find that a regularization strategy based on the most realistic prior information tends to offer the most reliable discrimination of facies. Our results confirm the ability of ERT surveys, when properly designed, to detect facies variations in alluvial aquifers. The method can be easily extended to other contexts.

Keywords: electrical resistivity tomography, alluvial aquifers, probability distribution, sensitivity, regularization, inversion, facies discrimination.

INTRODUCTION

Over the past two decades, electrical resistivity tomography (ERT) has become a standard geophysical technique in the field of hydrogeology (Binley et al., 2015). However, a relatively high level of uncertainty always accompanies the interpretation of the corresponding tomograms, which often renders the direct use of ERT images difficult. In particular, the interpretation of complex and heterogeneous geological systems such as alluvial aquifers and the quantitative integration of geophysical data into subsurface hydrological models require consideration of a number of important issues such as the petrophysical relationships linking geophysical and hydrological parameters (e.g., Doetsch et al., 2010), the spatially-dependent resolution of tomograms (e.g., Day-Lewis et al. 2005), and the effect of regularization on the resulting images (e.g., Caterina et al., 2014). To address these issues, several approaches have been developed including using synthetic simulations to explore petrophysical relationships (Moysey et al., 2005; Singha and Moysey, 2006), performing coupled hydrogeophysical inversion (Hinnel et al., 2011; Irving and Singha, 2010), using image appraisal tools to identify zones of the tomogram that can be reliably interpreted (Caterina et al. 2013, Beaujean et al. 2014), the use of novel regularization approaches (e.g., Blaschek et al., 2008; Hermans et al., 2012, Oware et al., 2013), and the incorporation of structural information into the inversion (e.g., Caterina et al., 2014; Doetsch et al., 2012a).

The above-mentioned studies have enhanced the imaging capabilities of ERT and broadened its range of applications, but they are generally not sufficient to answer the question of reliability of the results for a specific purpose. Indeed, in many cases, one wishes to perform a geological interpretation of the tomogram in terms of facies or hydrofacies. Given the limited resolution of the ERT experiment, the typically restricted number of ground-truth data, and the non-uniqueness of inverted geophysical models, carrying out such an interpretation is not straightforward. It is therefore critical that we develop methodologies to quantify the ability of

ERT to detect facies; that is, to indicate what facies is present at each subsurface location along with a corresponding uncertainty estimate.

To identify facies based on geophysical data, a number of studies have utilized multi-parameter analysis and clustering to perform subsurface zonation. Relating the clusters to facies through their geophysical parameters enables a straightforward classification. Fuzzy clustering, through partial memberships, also allows for uncertainty assessment. Paasche et al. (2006) proposed such a clustering approach to perform subsurface zonation based on georadar and seismic data. Paasche and Tronicke (2007) even included the clustering approach in the inversion procedure to obtain a common model of georadar and P-wave velocities with zonation. Doetsch et al. (2010) performed joint inversion of GPR, seismic and ERT data to obtain a three-facies zonation (gravel, sand, clay) in an alluvial aquifer. Although such clustering approaches are highly useful, a significant drawback is that they require having multi-method geophysical data, the latter of which are not always available.

In this paper, we focus on the ability of a single geophysical method (ERT) to provide a facies classification. To this end, we rely on a Bayesian framework for post-processing the outcomes of classical regularized inversion. The Bayesian framework offers an easy and straightforward way to express geophysical parameters in terms of categories, deriving a probability distribution for the sought parameters (here the facies or hydrofacies) instead of a single "best" estimate (Ezzedine et al., 1999; Rubin and Hubbard, 2005). The corresponding likelihood function can be estimated using co-located measurements of the inverted geophysical parameter and the textural description. Such a relationship enables accounting for the uncertainty resulting from the inversion process (regularization and noise) as well as data sensitivity. It is also straightforward to integrate such a Bayesian relationship into stochastic simulations involving Gaussian or multiple-point statistics. As an example, Ruggeri et al. (2013, 2014) developed a methodology to integrate geophysical data into hydraulic conductivity sequential simulations through a Bayesian relationship. They took into account the variable resolution of the tomograms by adapting the likelihood function to the level of uncertainty of their resistivity estimates. Hermans et al. (2015a) used a Bayesian framework to transpose ERT inversions into soft probability maps used to constrain multiple-point geostatistical simulations of an alluvial aquifer within the context of facies-based hydrogeological inversion. However, they did not integrate sensitivity dependence nor did they analyze the effect of regularization on the proposed probability maps.

Here, we assess the ability of ERT to identify facies in an alluvial aquifer. Alluvial aquifers are generally composed of several facies and lithologies (clay, loam, sand, gravel) with complex architectures and interconnections depending on the fluvial system (channels, bars, point bars, crevasse splays, floodplains, levees, etc.). In this context, borehole logs and classical hydrogeological tests (pumping, slug or tracer tests) may not be sufficient to capture the complexity of the deposits and determine their influence on groundwater flow and solute transport (e.g. Wildemeersch et al., 2014). ERT is a well-suited tool to study alluvial deposits as it is sensitive to the texture of the deposits (porosity, tortuosity, clay content) and to the pore-fluid. ERT has already been widely used to image deposits in alluvial aquifers and to improve the understanding of the depositional model (Baines et al., 2002; Bowling et al., 2005, 2007; Bersezio et al., 2007; Mastrocicco et al., 2010; Doetsch et al., 2010, 2012a). The use of ERT is also common in time-lapse mode to monitor changes taking place in alluvial deposits due to, for example, contaminant degradation (e.g., Chambers et al., 2010; Masy et al., 2016), salt tracer experiments (e.g., Doetsch et al., 2012b), or thermal tracer experiments (e.g., Hermans et al., 2015b).

The paper is organized as follows. First, the methods used in this study are described. Then, we analyze through a synthetic study the ability of ERT to discriminate between different facies. As ERT suffers from a loss of sensitivity with depth, low sensitivity regions are more affected by misclassification. We therefore propose to adapt our Bayesian framework to include the spatially varying sensitivity and counterbalance this effect. Next, we apply this learning to a field case, where we consider two different regularization procedures to investigate the role of the inversion method and its underlying assumptions on the different facies probability maps.

METHODS

ERT inversion

The ERT forward problem is non-linear and non-unique, typically characterized by a large number of subsurface model parameters and relatively few data (e.g., Aster et al., 2005). To address the issue of non-uniqueness in the corresponding inverse problem, regularization is normally used (e.g., Tikhonov and Arsenin, 1977) whereby prior information regarding the model parameters is considered in order to obtain a single solution. The basic idea of deterministic inversion is to minimize an objective function of the form

$$\psi(\mathbf{m}) = \psi_{\mathrm{d}}(\mathbf{m}) + \lambda \psi_{\mathrm{m}}(\mathbf{m}) \qquad (1)$$

where the first term on the right-hand side of equation (1) expresses the data misfit and the second term quantifies some undesired characteristic of the model, for example its “roughness” as measured by a first or second derivative operator. The regularization parameter λ then controls the balance between minimizing these two terms in the inversion procedure. In this paper, we consider two regularization operators. The “smoothness constraint” regularization operator, used in the vast majority of geophysical inversions, is used for both the synthetic and field cases. For the field case only, we also consider geostatistical regularization (Hermans et al., 2012), which is based on a prior definition of the model parameter covariance matrix.

To invert ERT data in this study, we used the finite-element inversion code CRTomo (Kemna, 2000) which seeks to iteratively minimize equation (1) until the root-mean-square of the weighted data misfit, $\varepsilon_{RMS} = \sqrt{\frac{\psi_d(\mathbf{m})}{N}}$ with N representing the number of data, reaches a value of 1. The overall aim is to fit the data to its assumed level of error under the condition that the model functional is minimized. Parameter λ is optimized during each iteration through a line search to minimize the data misfit.

Data sensitivity

Data sensitivity is commonly used as an image appraisal tool for ERT inversions (e.g., Caterina et al., 2013). As with resolution, the sensitivity of a tomogram varies spatially, generally showing a decreasing trend with depth for surface arrays. In this study, we use the error-weighted cumulative sensitivity as an indicator of the quality of an inversion result, which is defined as (Kemna, 2000)

$$\mathbf{S} = \mathrm{diag}\left\{\mathbf{J}^{\mathrm{T}}\mathbf{W}_{\mathbf{d}}^{\mathrm{T}}\mathbf{W}_{\mathbf{d}}\mathbf{J}\right\}, \qquad (2)$$

where $\mathbf{W_d}$ is the data weighting matrix, $\mathbf{J}$ is the Jacobian matrix and $^{\mathrm{T}}$ denotes the transpose operator. Values of $\mathbf{S}$ depend on the distribution of resistivity in the subsurface and data errors. We normalize $\mathbf{S}$ by its maximum value so that its values are lower or equal to one and we express it on a logarithmic scale. A high value means a strong influence of the corresponding model parameter on the predicted data and is therefore a favoring factor for a good resolution. Hereafter, we will use the term "sensitivity" to denote this normalized cumulative sensitivity.

Bayesian facies probability

To discriminate hydrofacies based on inverted ERT images, we use co-located ERT and facies description data to first determine the distribution of inverted resistivity $f(\rho|A_i)$ corresponding to each facies A_i as well as the facies proportion $P(A_i)$. We also consider sensitivity-dependent resistivity distributions $f(\rho|A_i,S_j)$, where S_j denotes model cells having a given sensitivity class. The latter is defined on a logarithmic scale by decade (i.e., from 1 to 10^{-1}, 10^{-1} to 10^{-2}, etc.). For the field case, co-located data correspond to borehole locations, where observations of facies can be compared with the ERT-derived resistivities. For the synthetic case, we have access to the facies distribution at all locations throughout the model space. Using Bayes' rule, we can then compute the conditional probability of observing a facies given the resistivity and sensitivity values as follows:

$$P(A_i \mid \rho, S_j) = \frac{f(\rho \mid A_i, S_j) P(A_i)}{\sum_i f(\rho \mid A_i, S_j) P(A_i)}. \quad (3)$$

Performance indicators

Equation (3) can be applied to an inverted resistivity profile and the resulting facies probabilities can be used to assess the ability of ERT to identify those facies. In this regard, we define two indicators of performance in this paper:

1) The classification performance (expressed in %). Each cell of the tomogram is assigned the facies having the highest probability and the number of correctly classified cells is counted.
2) The probability performance (expressed in %), which is calculated using

$$\xi = 100 \times \frac{1}{n} \sum_{k=1}^{n} \omega_k \max_i \left\{ P(A_i \mid \rho, S) \right\} \quad (4)$$

where n is the number of cells in the tomogram and ω_k is a weight equal to 1 if the cell is correctly classified according to the classification performance indicator, i.e. when the correct facies has the highest probability, and -1 otherwise. The philosophy of this indicator is to reward when a facies is correctly identified as the most probable and to penalize otherwise. In contrast with the classification performance indicator, it takes into account, through the maximum conditional probability $P(A_i|\rho,S)$, the uncertainty of the estimate. The final score is between -100 and +100%, the latter of which is obtained if all cells have a probability of 100% for the correct facies. A negative value indicates that the method is not appropriate to detect facies since it would provide in average more bad than correct information.

STUDY SITE

The study site for our field case is located in the alluvial aquifer of the Meuse River, in Hermalle-sous-Argenteau (Belgium) near the Dutch-Belgian border (Figure 1A and B), between the Meuse River and the Albert Canal. This is an experimental site of the University of Liege used for hydrogeophysical field experiments (e.g., Wildemeersch et al., 2014; Hermans et al., 2015b). A detailed description can be found in Hermans et al. (2015a, 2015b).

Alluvial deposits at the site are 10-m thick and mostly composed of sandy gravel with large zones of clean gravel (and pebbles) having a higher hydraulic conductivity and zones composed of loam, clay and clayey gravel of lower hydraulic conductivity. A total of 23 boreholes are available at the site with a corresponding textural description (Figure 1C and 2). Based on the observed heterogeneity, we chose to describe the deposits using three hydrofacies. The first facies is composed of low hydraulic conductivity deposits (clay and loam). The second facies is composed of clean gravel with high hydraulic conductivity. Finally, the third facies is made up of sand/sandy gravel and has an intermediate hydraulic conductivity. The respective facies proportions from borehole logs are 18%, 42% and 40%.

Clay deposits are mostly limited to the surface, and gravel is generally found below sand in the bottom part of the aquifer. The thickness of the sand is variable (Figure 2).

We conducted 12 parallel ERT profiles at the site (Figure 1C) using 64 electrodes spaced every 2 m (total length of 126 m) except for the northern profile which is only 102 m long. We used a dipole-dipole configuration to collect the data with dipole lengths from 2 to 18 m and dipole separations from 1 to 6 dipole lengths. We used reciprocal measurements to assess the quality of the data. A linear error model was deduced to weight the data during the inversion ($\mathbf{W_d}$ in equation 2) having the form (Slater et al., 2000)

$$|e| = a + b\mathrm{R} \tag{5}$$

where |e| is the absolute reciprocal error (in Ohm), *a* is the absolute error equal to 0.002 Ohm, *b* is the relative error equal to 0.26%, and R is the mean resistance (in Ohm).

SYNTHETIC VALIDATION

In this section, we use synthetic simulations to assess the ability of ERT to discriminate facies and we examine the effect of the varying data sensitivity on the classification and probability performance metrics. In the next section, a similar approach is applied to the Meuse River field data.

Description of the set-up

The synthetic benchmark is based on 96 synthetic models simulating alluvial deposits as they could be expected at the study site described in Section 3. The considered model domain is 126 m wide by 10 m deep, which was discretized into 1 m x 0.5 m cells to yield a total of 2540 model parameters. Below 10 m, we simulated the presence of resistive (300 Ohm.m) bedrock, which is not considered in our interpretations.

The alluvial deposits are composed of three hydrofacies corresponding to those observed at the field site. The sand facies is considered as the background and was prescribed a resistivity of 250 Ohm.m. The clay (100 Ohm.m) and gravel (140 Ohm.m) facies were simulated respectively as lenses and channels of various sizes using 3D Boolean simulations (Maharaja, 2008) with proportions similar to the ones observed at the field site. The resistivity values chosen for each facies were based on values observed in the field. A number of 2D sections were randomly selected from the 3D models (Figure 3A and D).

A dipole-dipole ERT data set was then simulated for each model and contaminated by noise having a similar level to that observed in the field. The data sets were subsequently inverted using the standard smoothness constraint. Facies probabilities were estimated after inversion in order to determine the classification and performance indicators for each model, both with and without taking the data sensitivity into account.

Inversion of synthetic models and sensitivity dependence

Figure 3 shows two examples of synthetic facies models along with their respective inverted resistivity tomograms and sensitivity distributions. It can be seen from the resistivity tomograms (Figure 3A-B and 3D-E) that ERT is able to qualitatively locate overall the clay and gravel facies, which correspond to low resistivity values inside the sand facies. However, smaller clay lenses are not detected and facies discrimination in the bottom part of the images is more challenging. Both of these effects are expected; the first is related to the resolution capabilities of ERT, whereas the second is related to the decrease of sensitivity with depth, as clearly illustrated by Figure 3C and F.

The presence of three different facies is also challenging in this case. Clay lenses and gravel channels both represent low resistivity anomalies compared to the background sand facies. Therefore, their juxtaposition among the deposits is difficult to resolve. More precisely, clay lenses in the bottom part of the deposits tend to display resistivity values close to those of the gravel facies.

The synthetic benchmark allows the comparison of the inverted resistivity value and facies identity for about 250,000 cells, thus providing a good test of the ability of ERT to detect facies as well as allowing us to analyze the dependency of the results upon data sensitivity. Figure 4A shows the mean and 95% interval of inverted resistivity for each facies as a function of the value of relative cumulative sensitivity, as well as the global values for the case where sensitivity is not taken into account (Total). We see that for high sensitivity values above 10^{-3}, which correspond to a maximum depth of about 5 m (Figure 3), the facies resistivities are well separated. In particular, sand and gravel show no overlap, meaning that ERT is able to almost unequivocally identify the presence of sand in high sensitivity regions (Figure 4B). It is more difficult to discriminate between clay and gravel due to their close resistivity values. For sensitivities between 10^{-3} and 10^{-5}, the sensitivity-dependent distributions are relatively close to the global distributions, although with smaller ranges of variation because they are not affected by the shallow cells like the total distribution. As a result, the corresponding conditional probability distributions computed from equation 3 (Figure 4B) are flatter and thus less discriminating. For sensitivity values below 10^{-5}, the sensitivity-dependent resistivity distributions are superimposed upon one another, which means that the resolving power of ERT for facies discrimination is weak or non-existent. Those sensitivity values are limited to the bottom corners of the grid or to the deep cells (bedrock) not considered here.

Overall, we see that the means of the sensitivity-dependent inverted resistivity distributions follow a clear trend towards a similar value when sensitivity decreases, which results from the decreasing resolution of ERT with depth. This value is related to the starting or reference model used for the inversion, which in this case is the mean apparent resistivity of the data.

Conversely, the 95% bounds do not follow a clear trend, but they are smaller than the ones from the global distributions.

The conditional probabilities (Figure 4B) can be used to transform each resistivity distribution into three probability maps, one for each considered facies. For a given cell, the three probabilities sum to 1. We illustrate the process in Figure 5 for the second model presented in Figure 3D, where we show the probability maps for gravel, sand, and clay that were obtained using the global probability distributions (Figures 5A to 5C) as well as the sensitivity-dependent distributions (Figures 5D to F). Gravel is the most challenging facies in our example since it corresponds to intermediate resistivity values, and will therefore never reach a probability equal to 1. We see in Figure 5 that high resistivities correspond to high probability values for sand whereas low resistivities correspond to high probability values for clay. In contrast, intermediate resistivities (around 160 Ohm.m) signify a higher probability for the gravel facies. Using sensitivity-dependent distributions, the values of the probability maps are more discriminant near the surface (first 5 m). Indeed, low values tend to be closer to zero and high values tend to be slightly closer to 1. This trend can be observed for all facies. Below 5 m, fewer changes are visible. We can detect a few locations where the probability of gravel, although relatively high, decreases when the sensitivity-dependent distributions are considered. The reason for this is that in low sensitivity regions, the sensitivity-dependent distributions are less discriminant than the global ones (e.g. Figure 5D, X = 30m and Y = 9m), which means that even if the inverted resistivity seems to correspond to gravel, we are less confident in this estimation and therefore attribute a lower probability to it. In turn, the probability of sand or clay increases.

Note that we do not a see strong differences in the probability maps when taking into account the sensitivity-dependent distributions.. The reason for this is that the design of the survey is such that the sensitivity values at 10 m depth are still reasonable (around 10^{-4}, see Figures 3C and F). For those sensitivity values, the sensitivity-dependent distributions are still at least as discriminant as the global ones (Figure 4). If we tried to interpret models below 10 m depth, we would observe a clear decrease in the probability. This illustrates the crucial importance of survey design if the sensitivity is not taken into account during interpretation. The comparison of global and sensitivity-dependent results with the probabilistic methodology for synthetic models can be used to assess the validity of survey design choices and estimate the depth of investigation; as long as the probability maps from the sensitivity-dependent relationships are more discriminant than the global distributions, the sensitivity can be considered as sufficient for interpretation. Global distributions will lead to conservative interpretations. For lower sensitivities, ignoring the sensitivity-dependence bears the risk of over-interpretation of the ERT images.

Facies discrimination performance

The more discriminant behavior of the sensitivity-dependent probabilistic relationship near the surface indicates that it will provide sharper probability estimates and better facies discrimination. We now quantify the benefit of using the sensitivity-dependent relationships with the performance indicators of Section 2.4.

For facies classification, the facies with the highest probability is assigned to each cell. An example is shown in Figure 6 for the model corresponding to Figure 3D. We use transparency to indicate the probability corresponding to the most probable facies. That is a fully transparent (white) cell corresponds to a maximum probability of 33%, i.e. for which all of the facies have the same probability and the classification is impossible. A completely opaque cell, on the other hand, corresponds to a facies probability of 100%. In comparison with the true facies (Figure 6A), we see that the classification of the sand facies is relatively accurate (Figure 6C), especially in the upper part of the model. Figure 6D is an indicator map of misclassification. Most misclassifications occur between the gravel and clay facies. This is expected given their respective resistivity. In the case of only two facies, ERT would therefore be more discriminating (Table 1).

Note in Figures 6E and 6F that there are few differences in the classification when sensitivity is taken into account. Globally, the same structures are observed. Even though sensitivity plays a role in the discrimination of facies, it remains the inverted resistivity value which orientates the classification. Indeed, with sensitivity dependence, a facies may have a higher or lower probability, but it may not always be sufficient to modify the order between facies and thus the classification.

The overall performance of the classification procedure is summarized in Table 1. On average for all of the models, 65% of the cells are correctly classified whatever method is used. The average risk of misclassification is therefore 35%. Within the 96 synthetic models, the minimum classification performance is about 50% and the maximum is 79%. As observed in Figure 6, misclassified cells are mainly between clay and gravel. If such misclassifications are disregarded, the average performance increases up to 79%. This means that increasing the number of facies to detect will inevitably degrade the performance. In our case, this is however mandatory because clay and gravel have opposite hydrogeological behaviors. Other misclassifications occur at the transition between facies.

Table 2 examines in greater detail the ability to classify cells according to sensitivity. Below values of 10^{-5}, less than 50% of the cells are correctly classified. Those cells generally correspond to low Bayesian probabilities, close to the prior probability, meaning that classification should be avoided because uncertainty is high. Interestingly, the sensitivity-dependent approach performs better at each sensitivity level, but the greatest improvement is observed for very low sensitivity values. At those levels, the sensitivity-dependent approach is more conservative and avoids cell misclassification. Misclassification above the bedrock also occurs because the resistivity at this depth is influenced by the more resistive bedrock, favoring classification into the more resistive sand facies. This aspect is discussed later in our field application.

The use of the transparency scale in Figure 6 indicates how the probability can be used to gain important information regarding the uncertainty of the classification. We see that, globally, classification into the sand facies is linked to a low uncertainty (opaque value), whereas classification into the gravel facies is less certain. Therefore, a misclassification within the gravel facies is less prejudicial because it is informed by a lower probability.

The transparency as an uncertainty estimate is directly linked to the probability performance indicator. Using the sensitivity-dependent approach, the probability performance indicator (Table 1) increases more significantly than the classification performance indicator. With sensitivity dependency, high-sensitivity correctly classified cells will contribute to higher rewards, and low-sensitivity poorly classified cells to less penalization. In contrast, high-sensitivity poorly classified cells will contribute to higher penalization, and low-sensitivity correctly classified cells to less reward. In the upper 5 m of the subsurface in Figure 6F, many correctly classified cells have a higher probability for the most probable facies compared to Figure 6D. This explains the increase of the probability indicator for the sensitivity-dependent case.

The probability performance metrics show that an increase in performance is observed when sensitivity-dependent distributions are used. The average performance increases from 24.5% to 27%, which corresponds well with the mean relative performance increase of 10%. The increase in performance is observed for 92 of the 96 tested models. In this specific case, few cells are located in very low sensitivity regions. For less favorable cases, the increase in performance could be more significant due to less penalization.

Our results show that the probabilistic approach, without sensitivity-dependence, is already relatively efficient when an ERT survey and corresponding goals are well-designed; that is, if we do not attempt to classify cells that have too low sensitivities or are located too deep. If only limited depths are considered, the global probabilistic framework can be considered as conservative.

FIELD RESULTS

We now apply the proposed methodology to field measurements. To this end, we consider one of the 12 ERT profiles acquired at the Meuse River site. We also demonstrate the effect of regularization on the resulting facies probability maps, as prior assumptions made during the ERT inversion have a large effect on the inverted resistivity distribution and, therefore, are expected to influence the facies probabilities.

Effect of sensitivity

Working with a field case is more difficult than our synthetic benchmark because we do not have access to many co-located measurements. Indeed, deriving directly sensitivity-dependent distributions from co-located measurements would require an unrealistically high number of boreholes in order to have reliable distributions. To overcome this limitation, we tried to reproduce the behavior observed for the synthetic cases in order to build a sensitivity-dependent relationship (Figure 7). To this end, we assumed linear trends for the mean of the distribution towards a common value at low sensitivity and 95% bounds equal to 80% of the ones observed for the global distribution. We acknowledge that this assumption is debatable, but it is made to propose a relatively straightforward way to apply the method in field conditions.

The large range observed for clay in the field is related to the geometry of the deposits at the Meuse River site. A layer of clayey loam is seen in almost all the boreholes; however, its thickness is small at some locations (0.5 to 1 m) compared to the electrode spacing (2 m), which makes it rather difficult to image. Inversion tends to show higher resistivity at these locations due to the presence of sand below the clay. In this section, we focus on the interpretation of the gravel facies, which is the most challenging to image given its intermediate resistivity.

The application of the methodology to the field case shows a behavior similar to the one observed for synthetic cases (Figure 8). Since the sensitivity values observed for the alluvial deposits are greater than 10^{-5} (Figure 8B), there is everywhere a tendency towards a more discriminant behavior of ERT when the sensitivity-dependency is used (Figure 8C and D). Zones of intermediate resistivity values see their probability of gravel increasing from 70% to approximately 80%, and a number of cells near the surface and near the bedrock have their probability decreasing to almost zero. This signifies that, if used as conditioning data for a hydrogeological model, ERT would give a stronger constraint on the geometry of the deposits. However, the global distributions already identify most trends in the deposits. Given the uncertainty in the chosen sensitivity-dependent relationship, we may conclude that, as was the case for our synthetic study, ignoring the sensitivity-dependence is a conservative approach because the choice of electrode spacing is sufficient to image the deposits down to a depth of about 10 m.

Effect of regularization

The probability maps obtained in Figure 8C and 8D are direct transforms of the resistivity values that were obtained through deterministic inversion of the ERT data. The solution of this inverse problem is non-unique and depends on the assumptions made about the subsurface resistivity distribution, which are expressed through the second term of the right hand side of equation 1. Here, we test the effect of changing the regularization operator for ERT inversion on the probability maps. To this end, we consider geostatistical regularization which uses the model parameter covariance matrix instead of a roughness matrix as the model constraint (Hermans et al., 2012; Caterina et al., 2014). To estimate the covariance matrix, we computed the variogram from borehole electromagnetic logs and found that a Gaussian model having vertical and horizontal correlation lengths of 4.4 m and 11 m, respectively, offered an acceptable fit (Hermans et al., 2015a). In addition, the known position (10 m depth) of the bedrock was imposed during the inversion. For the smoothness constraint regularization, a ratio of 2.5 between horizontal and vertical smoothing was used, which represents the same amount of anisotropy as the geostatistical constraint. Our objective is not to demonstrate that one inversion method is better than the other, which is clearly site or case specific, but rather to illustrate how regularization can modify our interpretation of the results. Therefore, the use of field data is more appropriate for this purpose.

The inverted resistivity distributions obtained using the two regularization methods are quite different, except near the surface (Figure 9A and B) where the inversion is mainly influenced by the data and not by the regularization operator. With the geostatistical regularization, the thickness of more resistive zones is limited and the decrease in resistivity corresponding to the presence of gravel above the bedrock is more pronounced. Globally, geostatistical regularization reproduces more satisfactorily the resistivity distribution measured with the electromagnetic log (Figure 10) in this specific case. This confirms how the incorporation of appropriate prior information into the inversion can improve the reconstruction of the resistivity distribution.

The different inverted resistivity values yield different resistivity distributions for each facies as well. The histograms obtained (based on the 12 ERT profiles) exhibit clear differences

(Figure 9C and D). For the smoothness constraint, the distributions of sand and gravel are close to each other, which makes the discrimination between these facies more difficult. With the geostatistical regularization, the mean value of the gravel facies is smaller. In both cases, the distribution for the clay facies is rather similar, because clay is mainly limited to the surface where both inversions yield similar results.

Using the corresponding histograms, the probability maps of gravel were computed (Figure 9E and F). A side-effect of regularization when attempting to identify three facies is that the transition between the low and high resistivity facies will almost always create a zone where the probability of the intermediate facies is high. This appears for example between X=80 m and X=120 m at 1.5 m depth. The patterns of high and low probability of gravel are relatively similar. The differences in the resistivity tomograms are partly counterbalanced by the probabilistic approach because they are taken into account in the conditional probability relationship, through the use of the co-located measurements histograms. However, some differences remain visible in the shape, amplitude, and position of some low-probability zones. In particular, the low probability of gravel observed at X = 60 m for the smoothness constraint is not in accordance with borehole data (Pz3, Figure 2). This results from the smoothness constraint, which is not able to image the decrease in resistivity at this location due to the presence of the underlying resistive bedrock, is a side-effect of using only three facies. Since the bedrock is not considered in our probabilistic analysis, this zone has a high probability of sand and would be deterministically classified as sand, with some uncertainty. However, if the bedrock were considered, this zone of low gravel probability would have a high probability of belonging to the bedrock, thereby decreasing the probability of sand. In this case, the misclassification is related to the uncertainty of the bedrock depth. This inconvenience is avoided in the geostatistical regularization because the position of the bedrock, known from boreholes, has been fixed during the inversion. The inversion process therefore identified that alluvial deposits could display a lower resistivity for which the most probable facies is gravel.

It is thus preferable to use the inversion method which gives the best estimate of the resistivity. In this case, the incorporation of prior information is helpful to discriminate facies more reliably. However, the choice of the regularization method is not related to the decision of using a probabilistic framework to interpret ERT images. Although the geostatistical regularization identifies a zone where the highest probability is for the gravel facies, the probabilistic framework provides probabilities around 0.7. This shows that this zone is uncertain with resistivity values that could correspond to gravel, sand or even clay.

The comparison of Figures 8 (C and D) and 9 (E and F) illustrates that the effect of regularization may dominate the effect of sensitivity-dependence on the derived probabilistic facies estimate, if very-low sensitivity zones are not considered. Indeed, the sensitivity-dependence, in contrast to the regularization, impacts the probability values but not the spatial resistivity distribution.

DISCUSSION AND CONCLUSIONS

In this paper, we have assessed the ability of electrical resistivity tomography to discriminate facies in alluvial aquifers. We propose to use a probabilistic relationship based on co-located facies and resistivity measurements to derive the resistivity distribution for each facies. Then, Bayes' rule is used to estimate the conditional probability of observing a facies given the resistivity. Our methodology has the advantage of integrating the uncertainty related to data noise and the inversion method into the estimate since it compares observed facies with post-inversion results.

We applied the methodology on a synthetic benchmark to verify the ability of ERT to correctly classify alluvial deposits into facies. The performance is quite good (on average, 65% correct classification) given the challenging task of determining three facies. Indeed, when only two facies are considered, the performance increases to 80%. Those results depend on the chosen synthetic models and resistivity values assigned to each facies. It is therefore necessary to be as close as possible to the actual field case to derive interesting guidelines. From our examples, it clearly appears that the classification ability decreases rapidly with depth. It is thus preferable to keep the Bayesian probabilities, which give additional information on the uncertainty of the estimates and can be further used in stochastic simulations, rather than classification for interpretation purposes. This can be done by either using probability maps for the interpretation, or by using a transparency scale related to the probability of the most probable facies for classification. The latter solution has the advantage of being able to express the results on a single figure.

The decrease of resolution and sensitivity of surface-based ERT data with depth is a well-known effect; the results of inversion become less certain with depth, whereas close to the surface ERT is more discriminant. We used the synthetic benchmark to analyze the influence of sensitivity on the performance of ERT to detect facies by adapting the probabilistic framework in order to account for the sensitivity-dependence. This did not appear to improve classification performance but it did allow us to increase confidence in the results by attributing, on average, higher probabilities to the correct facies. However, the approach using the global distributions already performs relatively well when the survey is properly designed (i.e., electrode spacing adapted to the desired depth of investigation).

Although determining the sensitivity-dependence of the resistivity distribution based on limited measurements was a challenge in our field case, we modeled this dependence based on the synthetic benchmark. However, the probabilistic approach is not designed to counterbalance all errors related to the inversion method. In this particular example, an increase of resistivity due to the presence of the bedrock was observed in the bottom part of the alluvial deposits, leading to the identification of a high probability of sand instead of gravel. This effect was avoided by using an appropriate inversion method integrating prior information regarding the position of the bedrock and a geostatistical constraint. The resistivity distribution obtained was closer to the true resistivity measured by a logging device. The probabilistic framework allows to associate an uncertainty estimate to the presence of a given facies.

The proposed methodology, using a large number of synthetic models, could be used to assess the ability of ERT to image various features in different contexts and to analyze the influence of other effects known to modify the inverted resistivity distribution such as the electrode configuration (Dahlin and Zhou, 2004) or the regularization trade-off parameter (Audebert et al., 2014). It can also be directly applied in Bayesian framework where geophysical measurements are used to update the probability of a given properties, an obvious example being soft conditioning of facies-based multiple-point geostatistical simulations (Hermans et al., 2015a).

The methodology can also be used as an alternative, probabilistic framework to identify if the survey is correctly designed regarding the objectives of the study and estimate the depth of investigation: as long as the sensitivity-dependent probabilities are similar to, or more discriminant than, the global ones, the survey will not systematically overestimates the ability of ERT to detect specific features in the subsurface.

ACKNOWLEDGEMENTS

This work was performed by T. Hermans while on a fellowship from the Belgian American Educational Foundation and Wallonia-Brusssels International for post-doctoral studies at Stanford University. We are grateful to Alain Dassargues and the Hydrogeology and Environmental Engineering unit of the University of Liege for giving us access to the experimental site of Hermalle-sous-Argenteau. We also thank two anonymous reviewers and the associate editor for their helpful comments.

References

Aster R. C., Borchers B. and Thurber C. 2005. *Parameter estimation and inverse problems*, Elsevier Academic Press, Amsterdam.

Audebert M., Clément R., Touze-Foltz N., Günther T., Moreau S. and Duquennoi C. 2014. Time-lapse ERT interpretation methodology for leachate injection monitoring based on multiple inversions and a clustering strategy (MICS).

Baines D., Smith D.G., Froese D.G., Bauman P. and Nimeck G. 2002. Electrical resistivity ground imaging (ERGI) : a new tool for mapping the lithology and geometry of channel-belts and valley-fills. Sedimentology, 49(3), 441-449.

Beaujean J., Nguyen F., Kemna A., Antonsson A. and Engesgaard P. 2014. Calibration of seawater intrusion models : Inverse parameter estimation using surface electrical resistivity tomography and borehole data. Water Resources research, 50(8), 6828-6849.

Bersezio R., Giudici M. and Mele M. 2007. Combining sedimentological and geophysical data for high resolution 3D mapping of fluvial architectural elements in the Quaternary Po plain (Italy). Sedimentary Geology, 202, 230-248.

Binley A., Hubbard S.S., Huisman J.A., Revil A., Robinson D.A., Singha K. and Slater L.D.. 2015. The emergence of hydrogeophysics for improved understanding of subsurface processes over multiple scales. Water Resources Research, 51, 3837-3866, doi:10.1002/2015WR017016.

Blaschek R., Hördt A. and Kemna A. 2008. A new sensitivity-controlled focusing regularization scheme for the inversion of induced polarization based on the minimum gradient support. Geophysics 73(2), F45-F54.

Bowling J., Rodriguez A., Harry D. and Zheng C. 2005. Delineating alluvial aquifer heterogeneity using resistivity and GPR data. Ground Water, 43(6), 8980-903.

Bowling J., Harry D., Rodriguez A. and Zheng C. 2007. Integrated geophysical and geological investigation of a heterogeneous fluvial aquifer in Colombus Mississippi. Journal of Applied Geophysics, 65, 58-73.

Caterina D., Beaujean J., Robert T. and Nguyen F. 2013. A comparison study of image appraisal tools for electrical resistivity tomography. Near Surface Geophysics, 11, 639-657.

Caterina D., Hermans T. and Nguyen F. 2014. Case studies of incorporation of prior information in electrical resistivity tomography: comparison of different approaches. Near Surface Geophysics, 12, 451-465, doi: 10.3997/1873-0604.2013070.

Chambers J.E., Wilkinson P.B., Wealthall G.P., Loke M.H., Dearden R., Wilson R., Allen D. and Ogilvy R.D. 2010. Hydrogeophysical imaging of deposits heterogeneity and groundwater chemistry changes during DNAPL source zone bioremediation. Journal of contaminant hydrology, 118(1), 43-61.

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

Day-Lewis F.D., Singha K. and Binley A.M. 2005. Applying petrophysical models to radar travel time and electrical tomograms: resolution-dependent limitations. Water Resources research, 110, B08206. doi:10.1029/2004JB003569

Doetsch J., Linde N., Coscia I., Greenhalgh S.A. and Green A.G. 2010. Zonation for 3D Aquifer Characterization Based on Joint Inversions of Multimethod Crosshole Geophysical Data. Geophysics, 75(6), G53–G64, doi:10.1190/1.3496476.

Doetsch J., Linde N., Pessognelli M., Green A.G. and Günther T. 2012a. Constraining 3-D Electrical Resistance Tomography with GPR Reflection Data for Improved Aquifer Characterization. Journal of Applied Geophysics, 78, 68–76, doi: 10.1016/j.jappgeo.2011.04.008.

Doestch J., Linde N., Vogt T., Binley A. and Green A.G. 2012b. Imaging and quantifying salt-tracer transport in a riparian groundwater system by means of ED ERT monitoring. Geophysics, 77(5), B207-B218.

Ezzedine S., Rubin Y. and Chen J. 1999. Bayesian method for hydrogeological site characterization using borehole and geophysical survey data : Theory and application to the Lawrence Livermore National Superfund site. Water Resources Research, 35(9), 2671-2683.

Hermans T., Vandenbohede A., Lebbe L. and Nguyen F. 2012. Imaging Artificial Salt Water Infiltration Using Electrical Resistivity Tomography Constrained by Geostatistical Data. Journal of Hydrology, 438–439, 168–180. doi:10.1016/j.jhydrol.2012.03.021.

Hermans T., Nguyen F. and Caers J. 2015a. Uncertainty in training image-based inversion of hydraulic head data constrained to ERT data: workflow and case study. Water Resources Research, 51, 5332-5352, doi :10.1002/2014WR016460.

Hermans T., Wildemeersch S., Jamin P., Orban P., Brouyère S., Dassargues A. and Nguyen F. 2015b. Quantitative temperature monitoring of a heat tracing experiment using cross-borehole ERT. Geothermics, 53, 14-26, doi: 10.1016/j.geothermics.2014.03.013

Hinnel A.C., Ferre T.P.A., Vrugt J.A., Huisman J.A., Moysey S., Rings J. and Kowalsky M.

2011. Improved extraction of hydrologic information from geophysical data through coupled hydrogeophysical inversion. Water Resources Research, 46, W00D40.

Irving J. and Singha K. 2010. Stochastic Inversion of Tracer Test and Electrical Geophysical Data to Estimate Hydraulic Conductivities. Water Resources Research, 46, W11514.

Kemna A. (2000), Tomographic inversion of complex resistivity – theory and application. PhD thesis. Ruhr-University of Bochum.

Maharaja A. (2008), TiGenerator: Object-based training image generator. Computer & Geosciences, 34, 1753-1761.

Mastrocicco M., Vignoli G., Colombani N. and Zeid N.A. 2010. Surface electrical resistivity tomography and hydrogeological characterization to constrain groundwater flow modeling in an agricultural field site near ferrara (Italy). Environmental Earth Sciences, 61(2), 311-322.

Masy T., Caterina D., Tromme O., Lavigne B., Thonart P., Hiligsmann S. and Nhuyen F. 2016. Electrical resistivity tomography to monitor enhanced biodegradation of hydrocarbons with Rhodococcus erythopolis T902.1 at a pilot scale. Journal of Contaminant Hydrology, 184, 1-13.

Moysey S., Singha K. and Knight R. 2005. A framework for inferring field-scale rock physics relationships through numerical simulation. Geophysical Research Letters, 32,L08304.

Oware E.K., Moysey S. and Khan T. 2013. Physically based regularization of hydrogeophysical inverse problems for improved imaging of process-driven systems. Water Resources Research, 49, 6238–6247 (2013).

Paasche H. and Tronicke J. 2007. Cooperative inversion of 2D geophysical data sets: a zonal approach based on fuzzy c-means cluster analysis. Geophysics, 72(3), A35-A39.

Paasche H., Tronicke J., Holliger K., Green A.G. and Maurer H. 2006. Integration of diverse physical-property models: subsurface zonation and petrophysical parameter estimation based on fuzzy c-means cluster analyses. Geophysics, 71(3), H33-H44.

Rubin Y. and Hubbard S. 2005. Stochastic forward and inverse modelling. In Y. Rubin and S. Hubbard (Eds.), Hydrogeophysics, 487-511, Springer, Dordrecht, the Netherlands.

Ruggeri P., Irving J., Gloaguen E. and Holliger K. 2013. Regional-scale integration of multiresolution hydrological and geophysical data using a two-step Bayesian sequential simulation approach. Geophysical Journal International, 194(1), 289-303.

Ruggeri P., Gloaguen E., Lefebvre R., Irving J. and Holliger K. 2014. Integration of hydrological and geophysical data beyond the local scale : Application of Bayesian sequential simulation to field data from the Saint-Lambert-de-Lauzon site, Québec, Canada. Journal of Hydrology, 514, 271-280.

Singha K. and Moysey S. 2006. Accounting for spatially variable resolution in electrical resistivity tomography through field-scale rock-physics relations. Geophysics, 71(4), A25-A28.

Tikhonov A.N. and Arsenin V.A. 1977. Solution of ill-posed problems. Winston & Sons, Washington.

Wildemeersch S., Jamin P., Orban P., Hermans T., Nguyen F., Brouyère S. and Dassargues A. 2014. Coupling heat and chemical tracer experiments for the estimation of heat transfer parameters in shallow alluvial aquifers. Journal of contaminant hydrology, 169, 90-99, doi: 10.1016/j.jconhyd.2014.08.001

TABLES

Estimator	**Without sensitivity**	**Sensitivity-dependent**
Mean classification performance (%)	65.26	65.32
Minimum classification performance (%)	49.29	50.31
Maximum classification performance (%)	79.17	78.94
Mean classification performance 2 facies (%)	78.92	78.87
Minimum classification performance 2 facies (%)	63.82	63.90
Maximum classification performance 2 facies (%)	90.43	90.55
Mean probability performance (%)	24.52	27.01
Minimum probability performance (%)	4.49	8.24
Maximum probability performance (%)	47.78	51.17
Mean relative performance increase (%)		10.15

Table 1. Classification and probability performances. The term “2 facies” means that misclassification between clay and gravel was disregarded.

Sensitivity classes	Number of cells	Without sensitivity (%)	Sensitivity-dependent (%)
$0 > S > 10^{-1}$	22825	84.21	84.61
$10^{-1} > S > 10^{-2}$	31066	84.18	84.48
$10^{-2} > S > 10^{-3}$	49633	71.19	71.33
$10^{-3} > S > 10^{-4}$	79928	59.28	59.40
$10^{-4} > S > 10^{-5}$	54329	52.33	52.50
$10^{-5} > S > 10^{-6}$	5696	43.43	45.05
$10^{-6} > S$	363	37.74	51.79

Table 2. Classification performance according to data sensitivity.

FIGURES

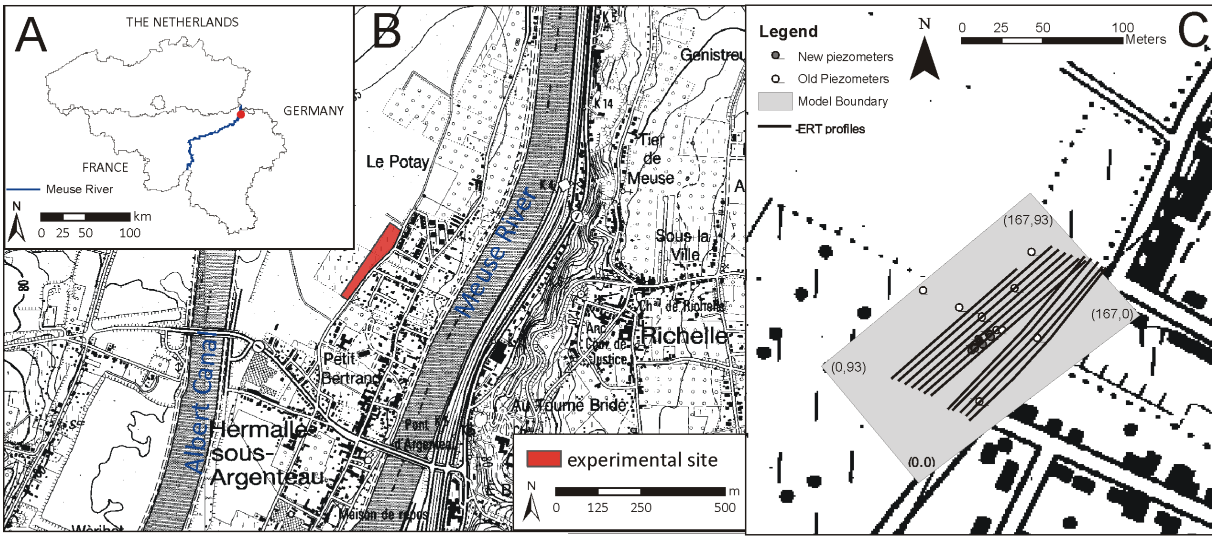

Figure 1. Location of the field site in the Meuse River alluvial aquifer (A and B) and of the boreholes used for facies description (C). The black lines show the position of 12 ERT profiles carried out on the site to study the resistivity distribution of the deposits.

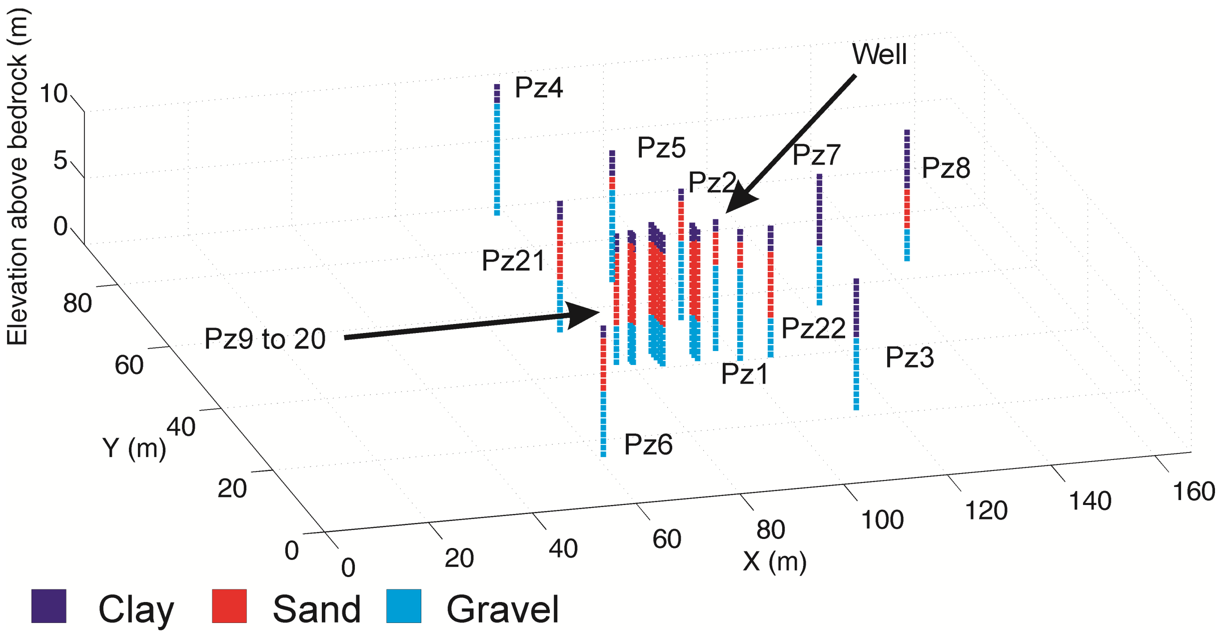

Figure 2. Interpretation of the borehole textural description data in terms of hydrofacies.

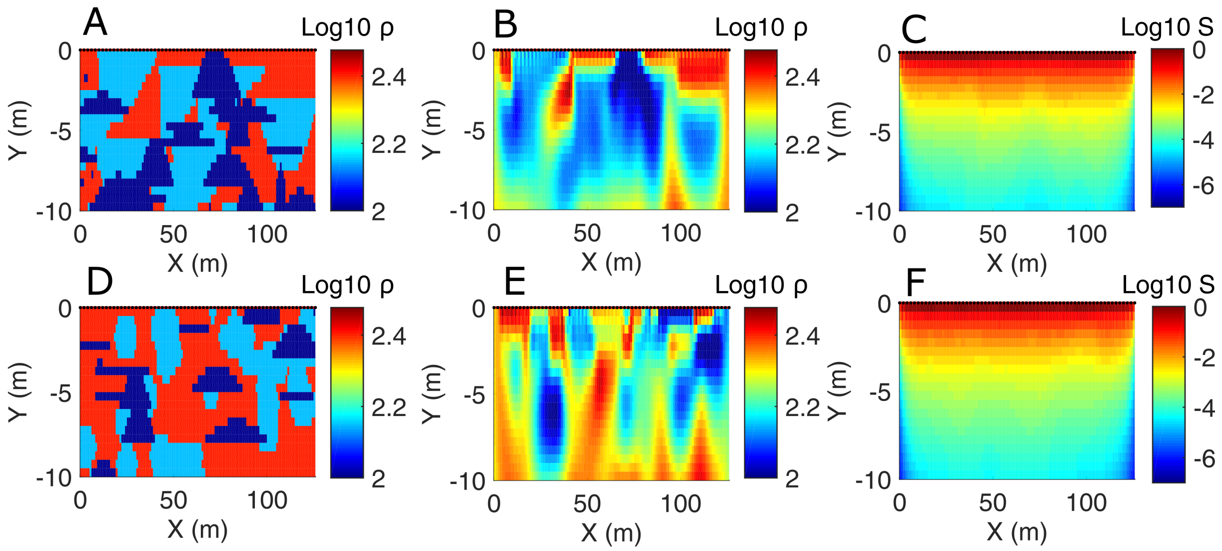

Figure 3. Two alluvial aquifer synthetic models (A and D) along with their respective inverted resistivity (B and E) and sensitivity distributions (C and F). The color of the resistivity of the facies is similar to the color scale used in Figure 2. The scale of the vertical axis is exaggerated.

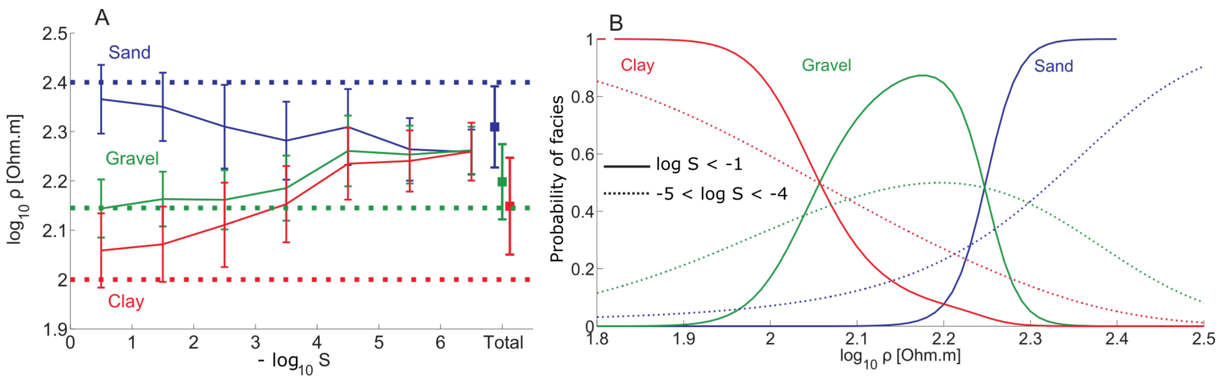

Figure 4. (A) Sensitivity-dependent resistivity distribution for the three considered facies (mean and 95% interval) for the synthetic benchmark, based on analysis of all 96 inverted models. The dotted lines correspond to the true value for each facies. (B) Conditional probability of the different facies given the inverted resistivity value for two different sensitivity classes, calculated using equation 3.

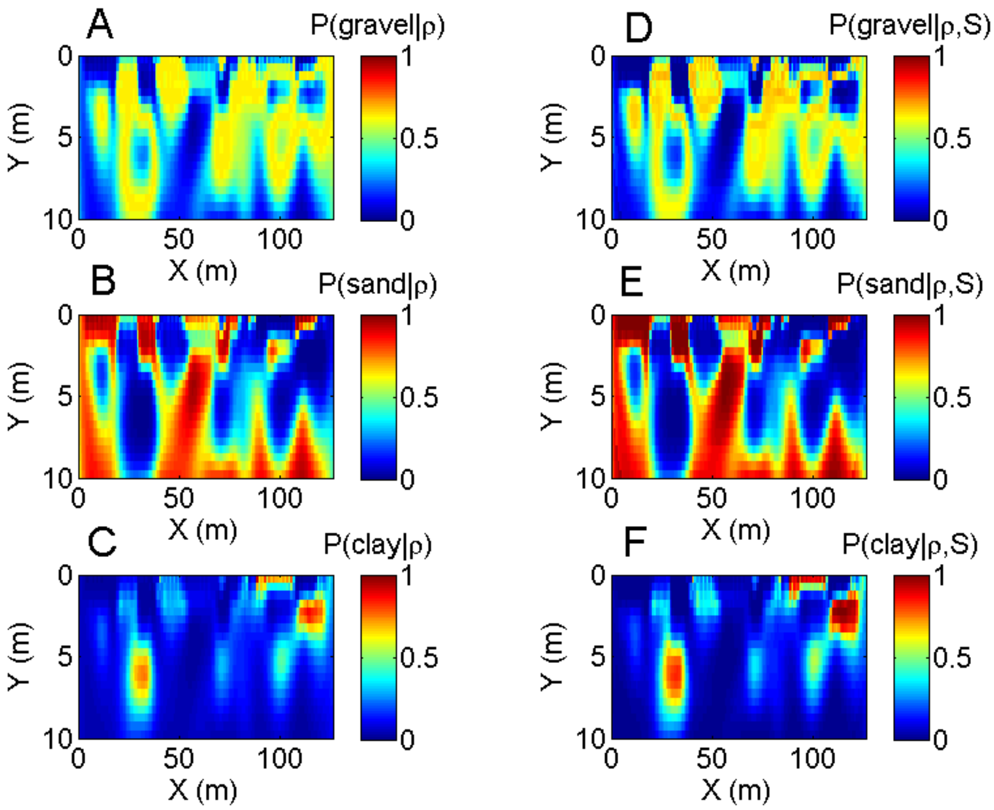

Figure 5. Probability of the three facies for the second model shown in Figure 3D, without (A, B and C) and with (D, E and F) the sensitivity-dependence taken into account. The scale of the vertical axis is exaggerated.

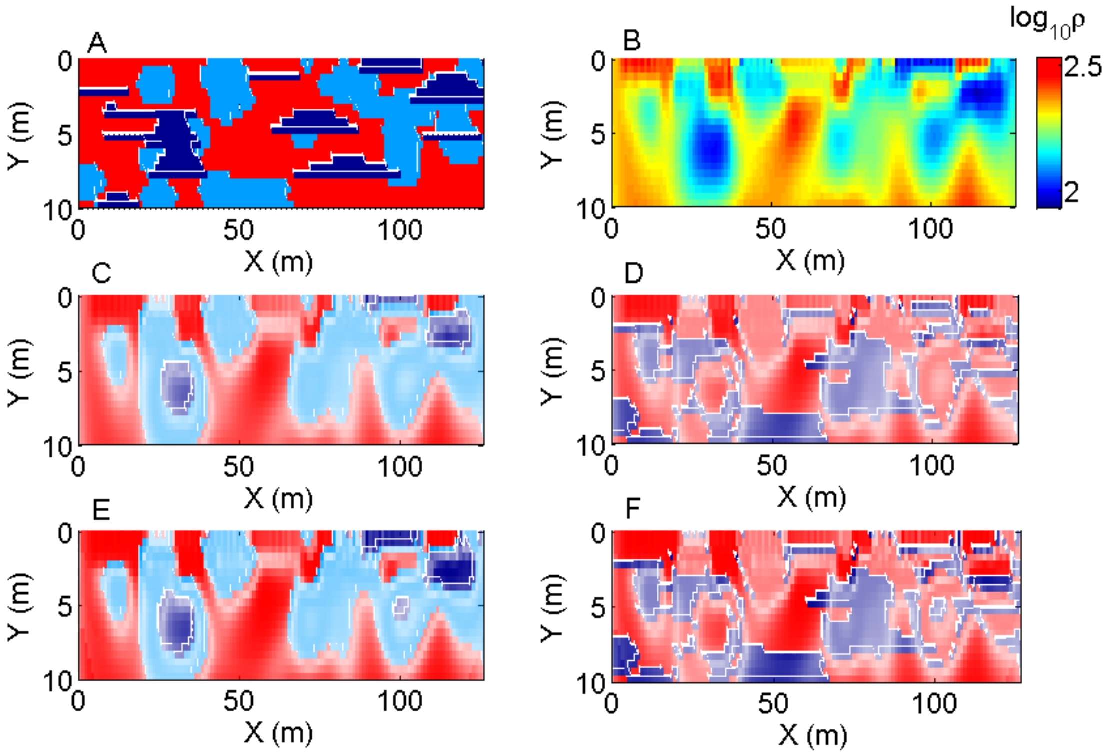

Figure 6. True model (A) from Figure 3D, inverted resistivity distribution (B), classifications based on the probability maps of Figure 5 without (C) and with (E) sensitivity dependence and corresponding misclassification indicator maps (D and F). Red is correct classification, blue corresponds to misclassification. The transparency scale in C-F is based on the probability of the most probable facies. Total opacity corresponds to a probability of 100%, total transparency to the minimum possible probability for the most probable facies: 33%. The color of the facies is similar to the color scale used in Figure 2. The scale of the vertical axis is exaggerated.

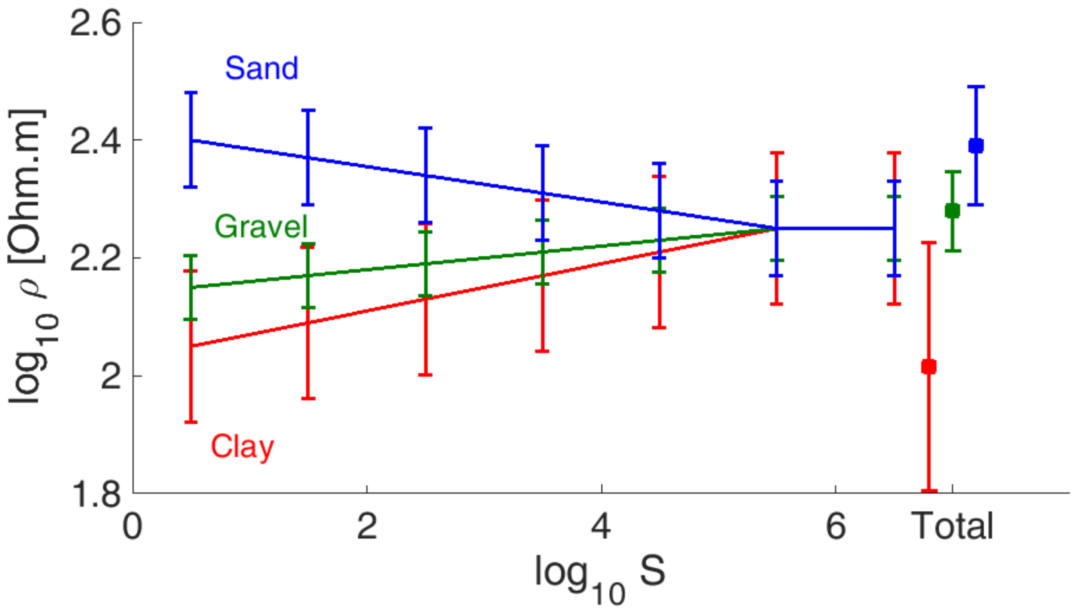

Figure 7. Assumed sensitivity-dependent resistivity distribution for the three considered facies (mean and 95% interval) for the field case.

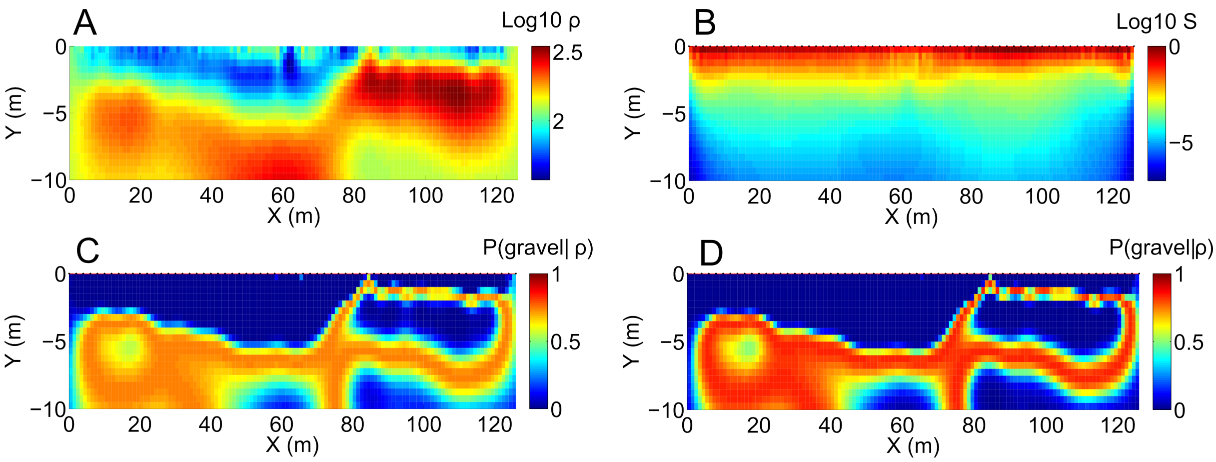

Figure 8. Smoothness constraint inversion of a field profile (A), the corresponding sensitivity distribution (B), and probability maps without (C) and with (D) taking sensitivity into account.

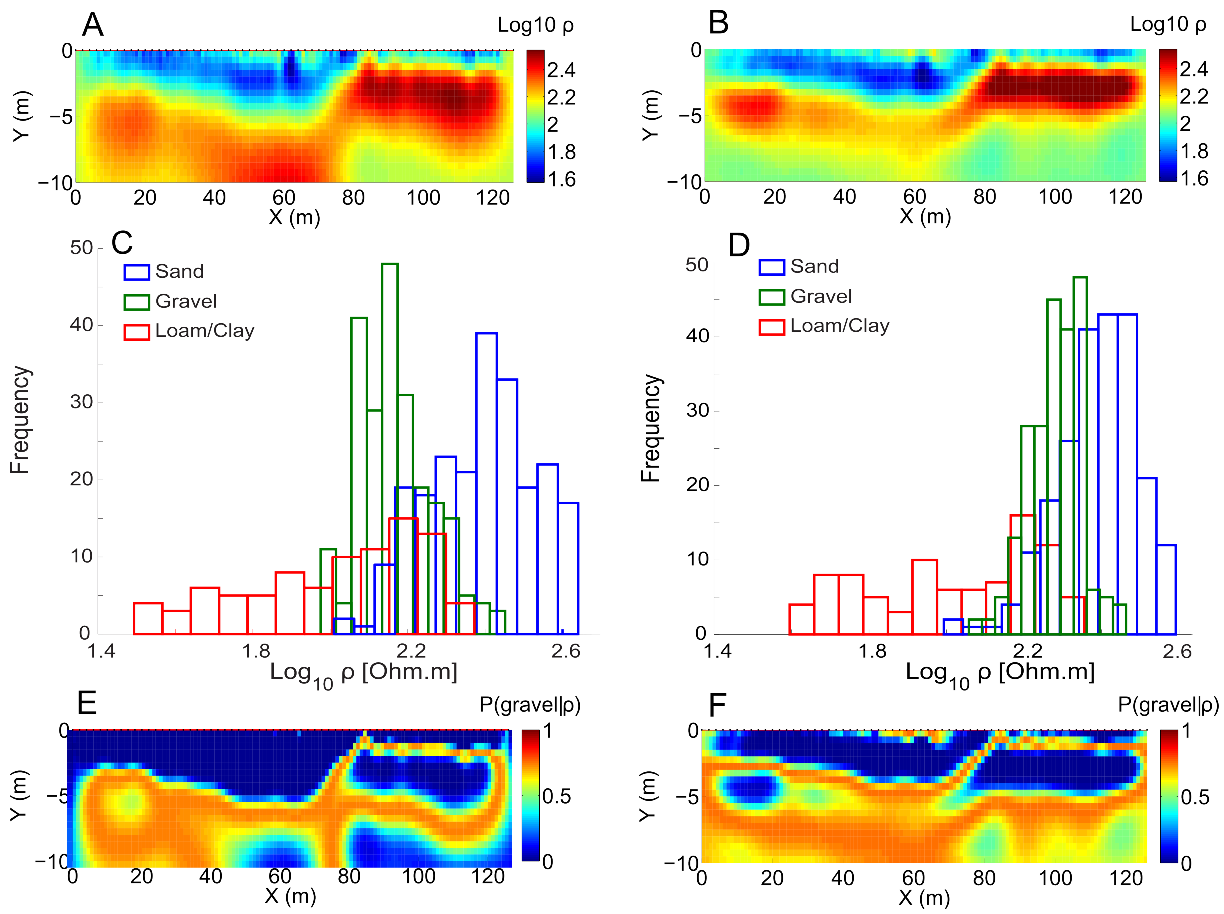

Figure 9. Effect of regularization on the gravel facies probability map. ERT inversions obtained using smoothness constraint inversion (A) and geostatistical inversion (B), respective histograms of resistivity for each facies (C and D) deduced from borehole logs, and corresponding probability map for gravel (E and F).

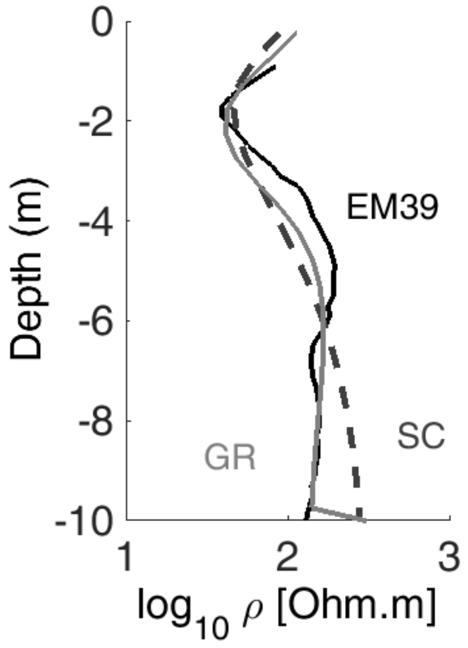

Figure 10. Comparison of inversion results with electromagnetic log in the middle of the profile of Figure 9 (GR = geostatistical regularization, SC = smoothness constraint). Geostatistical regularization is able to reproduce the decrease of resistivity above the bedrock.