

Case History

Quantitative interpretation of geoelectric inverted data with a robust probabilistic approach

Itzel Isunza Manrique¹, David Caterina¹, Frederic Nguyen¹, and Thomas Hermans²

ABSTRACT

The nonuniqueness of the solution to the geophysical inverse problem can lead to misinterpretation while characterizing the subsurface. To tackle this situation, ground-truth information from excavations and wells can be used to improve, calibrate, and interpret inverted models. We refer to quantitative interpretation as the decision analysis based on probability theory, which is focused on solving a classification problem. First, we present a probabilistic approach to classify the different types of materials or “categories” observed in borehole logs using multiple data sources: inverted 2D electrical resistivity tomography and induced polarization data and the positions (x , z) of these boreholes. Then, using the Bayes’ rule and permanence of ratios, we compute the joint conditional probabilities of each category, given all data sources in the whole inverted model domain. We validate this approach with synthetic data modeling for a complex

anthropogenic-geologic scenario and using real data from an old landfill. Afterward, we assess the performance of the probabilistic approach for classification and compare it with the machine learning algorithm of multilayer perceptron (MLP). In addition, we analyze the effect that the different data sources and the number of boreholes (and their distribution) have on both approaches with the synthetic case. Our results indicate that the MLP performance is better for delineating the different categories where the lateral contrasts in the synthetic resistivity model are small. Nevertheless, the classification obtained with the probabilistic approach using real data seems to provide a more geologically realistic distribution. We conclude that the probabilistic approach is robust for classifying categories when high spatial heterogeneity is expected and when ground-truth data are limited or not sparsely distributed. Finally, this approach can be easily extended to integrate multiple geophysical methods and does not require the optimization of hyperparameters as for MLP.

INTRODUCTION

Geophysical imaging has been widely used to derive insights into subsurface structures and processes by mapping physical parameters through noninvasive techniques. After processing and interpretation of geophysical measurements, this information can be translated into geologic structures, quantification of volumes, and geometries, or can be used to provide insight into groundwater processes (Romero-Ruiz et al., 2018; Whiteley et al., 2019; Slater and Binley, 2021). To obtain

these images or physical models, an inverse problem, which is usually ill-posed and presents a nonunique solution, is solved (Aster et al., 2018). Inverse problems may be solved using prior information in the form of structural or geostatistical constraints in the higher dimensional space of the model parameters (Kaipio et al., 1999; Chasseriau and Chouteau, 2003; Caterina et al., 2014) and more recently in lower dimensional spaces using machine learning methods provided that generative models may be trained to enforce consistent spatial patterns (Lopez-Alvis et al., 2021, 2022). Interpretation of

Manuscript received by the Editor 8 March 2022; revised manuscript received 14 December 2022; published ahead of production 23 January 2023; published online 20 April 2023.

¹University of Liege, Urban and Environmental Engineering, Liege, Belgium. E-mail: iisunza@uliege.be (corresponding author); david.caterina@uliege.be; f.nguyen@uliege.be.

²Ghent University, Department of Geology (WE13), Gent, Belgium. E-mail: thomas.hermans@ugent.be.

© 2023 The Authors. Published by the Society of Exploration Geophysicists. All article content, except where otherwise noted (including republished material), is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC). See <https://creativecommons.org/licenses/by-nc/4.0/>. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its digital object identifier (DOI). Commercial reuse is not permitted. The same license does not have to be used for derivative works.

geophysical images can be achieved with several approaches ranging from a qualitative analysis that might be validated by correlation with ground-truth data (Yannah et al., 2017; Magiera et al., 2019) to automated quantitative processes to directly assist interpretation when a large amount of data are available, for example, using deep learning approaches (Wang et al., 2018) or when the translation from bulk geophysical properties to properties of interest is sufficiently linear (e.g., Hermans et al., 2012). In the following, we focus on the quantitative approaches to improve data interpretation.

In geosciences, there is a recent increase in the use of machine learning, deeply rooted in applied statistics, in which computational models are built using inference and pattern recognition (Dramsch, 2020). In this context, supervised and unsupervised learning algorithms of linear and nonlinear methods also have been developed and adapted (Scheidt et al., 2018).

For example, Moghadas and Badorreck (2019) used neural networks to successfully link time-lapse electrical resistivity tomography (ERT) data to soil moisture, collecting reference data via an excavated pit and using the reference electrical conductivity and temperature values as data sources or inputs to train the supervised learning algorithm. To solve a classification problem also using supervised learning, Lysdahl et al. (2022) applied the algorithm of multilayer perceptron (MLP) to extract the depth to bedrock from airborne electromagnetics and sparse drillings. Training data were pairs of known depth points and resistivity data. The former approach was used in a field case with postglacial geomorphology, yet the authors concluded that the geologic complexity was the main limitation on the MLP performance.

Combining two or more geophysical methods based on different physical properties can greatly reduce the ambiguities inherent to each method (Hellman et al., 2017), improving the interpretation and geophysics-based characterization. Paasche et al. (2006) adopted a statistical approach to integrate the physical models from individually inverted georadar and seismic data into one multiparameter model and to estimate the spatial distribution of petrophysical parameters (from limited geophysical and petrophysical databases) using fuzzy *c*-means clustering. Another example of unsupervised learning to solve a classification problem is given by Whiteley et al. (2021), who use a Gaussian mixture model algorithm to classify geophysical data into cluster groups to build a ground model and characterize landslide materials. They use three geophysical variables as data sources: resistivity, P- and S-wave velocities, and a spatial variable: depth from the ground surface. In the context of landfill investigations, Inauen et al. (2020) apply several algorithms of supervised learning to classify geoelectric and seismic data according to the materials observed in several trial pits and boreholes. The main goal was to derive a model of a solid waste landfill, for which the algorithm of MLP presented a good classification performance.

In the literature, we find fewer applications for the interpretation of geophysical data based on statistics or probability theory exclusively. For example, Dewar and Knight (2020) develop a methodology for estimating the top of the saturation zone from airborne electromagnetic data (1D resistivity models) and measurements from nearby wells. The methodology included the optimization of two parameters: (1) a search radius to integrate resistivity data within this area and (2) statistical properties of the resistivity distribution that best captures the transition from an unsaturated to a saturated zone, i.e., the minimum, the maximum, the difference between the minimum and maximum, the mean, the difference between the 75th and the 25th percentiles, and the standard deviation.

In most of the cases discussed previously, the interpretation of geophysical data provides a single model representing the physical reality in which the uncertainty often is not considered. In this regard, the Bayesian framework has become one of the leading paradigms to quantify uncertainty in geophysical modeling, inversion, and data interpretation, which can be translated into more rigorous decision making in subsurface systems, especially under noisy data (Ray et al., 2018; Scheidt et al., 2018; Bobe et al., 2020; Parsekian et al., 2021). The Bayesian analysis or inference refers to all procedures that use the Bayes' theorem, where a quantitative relation is introduced to link predefined knowledge to new observations, thus comprising the computation of posterior distributions of a set of priors given a likelihood function (Piana Agostinetti and Bodin, 2018). For instance, Wellmann et al. (2018) combine the pre-existing geologic modeling with additional geologic considerations and gravity data simulations, applying the Markov chain Monte Carlo to evaluate and sample from the posterior distributions to obtain suitable geophysical models. This approach successfully addressed uncertainty and optimized the geologic model of a sandstone greenstone belt. More recently, Fossum et al. (2022) use the ensemble randomized maximum likelihood to update the subsurface uncertainty in earth models and simulate electromagnetic logs generated with generative adversarial networks and a forward deep neural network, respectively.

In line with the Bayesian framework, another approach for the interpretation of already inverted geophysical data was used by Hermans and Irving (2017), where the inverted parameters were expressed in terms of categories defined by the probability distributions of hydrofacies. They assess the use of ERT to identify and classify the hydrofacies in alluvial aquifers, using collocated inverted data and boreholes records, integrating the effect of the sensitivity spatial variation.

In this contribution, we use a probabilistic approach as an alternative to performing a rapid quantitative interpretation by classifying geophysical data according to the materials observed in a limited number of collocated borehole logs (hereafter referred to as categories or classes). The method is based on the previously mentioned approach used by Hermans and Irving (2017). We extend it to account for more than one geophysical model (i.e., ERT and time-domain induced polarization [IP]) and to include spatial trends in the collocated data. For comparison, we also apply the supervised machine learning algorithm of MLP and include the same data sources or input for training. We apply both approaches in a synthetic case of study and analyze the effect that the data sources and the number of boreholes (and their distribution) have on the probabilistic approach and MLP. Finally, we compare and validate the approaches using the geophysical data acquired in an old landfill with available collocated trial pits. Such systems are notably difficult to characterize with geophysics due to the strong heterogeneity and physical contrasts encountered and with boreholes due to the increased contamination risks. We observed that the probabilistic approach is suitable to classify inverted models that are highly variable along the whole domain, with predictions consistent with the prior sampling information and the integration of the prediction uncertainty. Note that a robust characterization is crucial in decision making for sustainable management, e.g., estimate the internal structure of a landfill to assess the potential for resource recovery and to prevent or evaluate the associated environmental pollution and prioritize (re)development scenarios (Jones et al., 2018; Van De Vijver et al., 2020, 2021). Next, we first present the methods used for the classification after inversion, the results are then presented for the synthetic case and the field data, followed by the "Discussion" and "Conclusion" sections.

METHODS OF CLASSIFICATION: PROBABILISTIC APPROACH AND MLP

Input data

To compare the probabilistic approach and MLP, we used the same input or training data. These are the collocated inverted resistivity ρ , chargeability C , the position (x, z) of the boreholes, and target categories. In these data, a relative cumulative sensitivity threshold is used to keep only parts of the tomograms (ρ and C) that are sufficiently well covered (e.g., [Beaujean et al., 2014](#)). Consequently, this also contributes to keeping reliable training data and reduces misclassification in the final model ([Hermans and Irving, 2017](#)).

Probabilistic approach

In the classification of the probabilistic approach, we first use the training data to define individual probability distributions, and then we compute the joint conditional probabilities of each category, given $\rho, C, x,$ and z in the whole domain of the inverted models. The results are given in terms of probability maps that belong to each predefined category which can be translated into classes. In the following sections, we first introduce the permanence of ratios, which is an alternative to computing joint conditional probabilities of different sources in the presence of data interdependence ([Journel, 2002](#)). Then, we describe the procedure of the probabilistic approach where the permanence of ratios is used.

Principle of permanence of ratios

Let A be an unknown event that can be assessed with two data events from different sources, B and D , through its conditional probability $P(A|B, D)$. For instance, A may represent a category, such as inert waste, and events B and D represent resistivity and chargeability. The easiest way to recombine these probabilities is to assume independence of the data events in which case the joint probability is the product of the marginal probabilities. However, this is a strong hypothesis as B and D are related to the common event A . To take this into account, [Journel \(2002\)](#) proposes an alternative to combine probabilities of different sources based on the permanence of updating ratios while guaranteeing all limit conditions (e.g., $P(A|B, D) \in [0, 1]$) even in the presence of complex data interdependence. The principle of permanence of ratios indicates that the rates or ratios of increments are typically more stable than the increments themselves. For simplicity, let us consider only two data events from different sources B and D , then the logistic-type ratio of the marginal probability of the unknown event A is

$$a = \frac{1 - P(A)}{P(A)} = \frac{P(\tilde{A})}{P(A)} \in [0, \infty], \tag{1}$$

where \tilde{A} is the complement of A . And similarly,

$$b = \frac{1 - P(A|B)}{P(A|B)} = \frac{P(\tilde{A}|B)}{P(A|B)}, \quad d = \frac{P(\tilde{A}|D)}{P(A|D)},$$

$$X = \frac{1 - P(A|B, D)}{P(A|B, D)} = \frac{P(\tilde{A}|B, D)}{P(A|B, D)} \geq 0. \tag{2}$$

Then, the ratio a can be seen as a measure of prior uncertainty about A : $a = 0$ if A is certain to occur and $a = \infty$ if A is an impos-

sible event. Similarly, d can be seen as the distance to A occurring after observing the data event D . The ratio d/a is then the contribution of D to that distance starting from the prior distance a . Finally, X would be the distance to A occurring after observing events B and D , and the ratio X/b is the incremental contribution of D starting from the distance b . The permanence of ratio assumes

$$\frac{X}{b} \cong \frac{d}{a}, \tag{3}$$

which means that the incremental contribution of data event D to the knowledge of A is the same after or before knowing B . Then, the joint conditional probability of the two events B and D can be expressed as $P(A|B, D) = 1/(1 + X) = a/(a + bd)$.

We can generalize the previous expression to n data events $G_i, i = 1, \dots, n$. Denoting by $|G_i, i = 1, \dots, n$ the joint conditioning to all n data events, the conditional probability provided by a succession of $(n - 1)$ permanence of ratios is

$$P(A|G_i, i = 1, \dots, n) = \frac{1}{1 + X} \in [0, 1], \tag{4}$$

with

$$X = \frac{\prod_{i=1}^n g_i}{a^{n-1}} \geq 0 \Leftrightarrow \text{Ln}(X) - \text{Ln}(a) = \sum_{i=1}^n [\text{Ln}(g_i) - \text{Ln}(a)], \tag{5}$$

$$a = \frac{1 - P(A)}{P(A)}, \quad g_i = \frac{1 - P(A|G_i)}{P(A|G_i)}, \quad i = 1, \dots, n, \tag{6}$$

which is an expression that verifies all limit properties, and it only requires the knowledge of the prior probability $P(A)$ and the n elementary single data event-conditioned probabilities $P(A|G_i)$, which can be evaluated independently one from another using collocated data.

The procedure of the probabilistic approach

Let A_i be the different categories (materials) found in the boreholes along the inverted sections. First, we estimate a prior probability or material proportion $P(A_i)$ based on the area they occupy in the boreholes collocated in the 2D inverted sections. Then, we determine unimodal distributions of the training data given each category A_i , i.e., $f(\rho|A_i), f(C|A_i), f(x|A_i),$ and $f(z|A_i)$. These distributions define the parameters, i.e., shape, locations, and scales, which we use afterward to estimate the distributions in an extended data set, i.e., $\rho, C, x,$ and z in the whole inverted model domain. In the next step, we compute the conditional probability of each material using the Bayes' rule. For instance, to compute the conditional probability of a category A_i , given the resistivity, we use

$$P(A_i|\rho) = \frac{f(\rho|A_i)P(A_i)}{\sum f(\rho|A_i)P(A_i)}, \tag{7}$$

and similarly, for the conditional probability of A_i given $C, x,$ or z .

At this stage, we need to combine the prior probabilities $P(A_i)$ with the conditional probabilities of A_i given the data sets into the

joint conditional probabilities $P(A_i|\rho, C, x, z)$. To this aim, we use equation 4 where the event A becomes the categories A_i , $n = 4$ and the data events G_i are ρ , C , x , and z . Therefore, the ratios $g_i = (1 - P(A|G_i))/P(A|G_i)$ will be given in terms of the marginal probabilities $P(A_i|\rho)$, $P(A_i|C)$, $P(A_i|x)$, and $P(A_i|z)$ for each category. The joint conditional probabilities are then normalized (divided by $\sum_i P(A_i|\rho, C, x, z)$) to ensure the closure condition, i.e., $P(A|\cdot) + P(\bar{A}|\cdot) = 1$. Then, the results are given in terms of probability maps for each category in the whole inverted model domain, and therefore may be used to assess the interpretation uncertainty. Finally, we also may derive a map in terms of the categories by comparing the normalized joint conditional probabilities of A_i and selecting the category corresponding to the largest probability value, i.e., classification model.

Supervised machine learning: MLP

Classification and regression methods are part of statistical learning and machine learning for which numerous methods have been developed, from simple linear regression to nonlinear methods such as neural networks or deep learning (Scheidt et al., 2018). Here, we focus on the multiclass classification problem, where we want to predict discrete class labels or categories for unlabeled patterns based on observations. We used the algorithm of MLP or a feedforward neural network, which proved to have a good performance for classification in the context of landfill investigations (Inauen et al., 2020).

Description of MLP

As explained by Goodfellow et al. (2016), the goal of MLP is to approximate some function f^* . For a classifier, the function $A_i = f^*(\{\cdot\})$ maps input data $\{\cdot\}$ to a category A_i . This algorithm defines a mapping $A_i = f(\{\cdot\}, \theta)$ and learns the value of the parameters θ (weight and bias coefficients of the transformation function) that result in the best approximation. MLP or feedforward neural networks are models where the information flows through the function from the input data, through the intermediate computations (linear and nonlinear data transformations followed by an activation function) used to define f and finally to the output A_i . They are networks because they can be composed of many different functions connected in a chain, i.e., $f(\{\cdot\}) = f^{(3)}(f^{(2)}(f^{(1)}(\{\cdot\})))$, where superscript 1 refers to the first layer of the network, 2 refers to the second layer, and the final layer is called the output layer. The overall length of the chain gives the depth of the model: the more layers the “deeper” the model. In this contribution, we are dealing with a small amount of geophysical data, thus as shown in the next sections, a simple neural network (few layers in the chain) proves to be enough.

The neural network makes use of training data to drive $f(\{\cdot\})$ to match f^* . These data provide approximations of f^* evaluated at different training points, which are accompanied by a category (A_i). Then, the learning algorithm decides how to use the other layers to produce the desired output, and as we cannot see the intermediate output of each of these layers, they are referred to as the hidden layers. Finally, each hidden layer of the network is composed of several units or neurons that can act in parallel representing a vector-to-scalar function.

In the multiclass classification, the output layer receives the values from the last hidden layer and transforms them into different classes, commonly with the Softmax function. This activation

function of the last layer normalizes the data and transforms them into an output probability vector, based on which the output classes are selected (e.g., Williams and Barber, 1998). Therefore, the output is quite similar to the probabilistic approach.

MLP architecture

To design and optimize the architecture of MLP in terms of the hyperparameters (such as the number of hidden layers, neurons, or regularization), the total training data are divided into a validation data set (10%–20%) and a remaining training data set (70%–80%). The partition is done randomly but preserves the relative frequency of the categories.

To tune the hyperparameters, we trained the MLP algorithm using combinations of different numbers of hidden layers (from 1 to 10), a number of neurons (1–100), a solver for weight optimization, and an activation function for the hidden layers computing the accuracy score or the fraction of correct predictions in the validation data set. Then, we select three MLP architectures from which the highest scores were obtained and compare them with the accuracy scores of the prediction in the training data set for several regularization values. We select the regularization parameter where the gap between the accuracies (from validation and training) is reduced while still preserving a relatively large accuracy score, i.e., generalization. In addition, as we have a multiple-classes problem, we first applied one-hot encoding to define each class (Potdar et al., 2017; Fu et al., 2019; Liu et al., 2021). This means we encoded the categories as a binary (one-hot) numeric array. Once the hyperparameters are selected and the architecture of the algorithm is defined, the training data set is the same as the one used for the probabilistic approach (or validation plus training data as indicated here).

Classification performance assessment

For both methods, we evaluate the performance of the classification or prediction of classes using a test data set. In the synthetic case, these are the categories known at the whole model space (excluding the data at the boreholes), and in the field case, a percentage of the training data where the categories are known. Then, we compute two classification scores to compare the results of this classification or “prediction” with the original categories of the test data set. First, we use the accuracy score, which computes the fraction of correct predictions. Given a predicted class \hat{A}_i for a sample i , the accuracy score can be expressed as

$$\text{accuracy}(A, \hat{A}) = \frac{1}{n} \sum_{i=0}^{n-1} 1(\hat{A}_i = A_i), \quad (8)$$

where A_i is the true category, n is the number of samples, and $1(x)$ is the indicator function, which maps the elements of a subset to one and the rest of the elements to zero (Pedregosa et al., 2011). We selected this accuracy score as we want to assess the classification performance of both methods when considering a model with a predominant material or class. In addition, we used the confusion matrix to assess the performance of the classification. In this matrix, the elements on the diagonal are the percentage of categories that are correctly predicted, and the off-diagonal elements are the misclassified percentage. In the synthetic and field case examples, the

categories for the classification are nonorganic waste deposits, soil, lime, backfill, and limestone bedrock.

RESULTS

Synthetic case study

Model generation and inversion

The synthetic model is inspired by a real near-surface scenario composed of anthropogenic materials deposited on a limestone quarry. The data were simulated for resistivity and chargeability distributions composed of five regions derived from nonorganic waste deposits, soil, lime, and backfill on a limestone quarry (Figure 1). In this geometry, we defined a surficial and continuous layer of waste, as this material was observed on the ground surface in the real landfill. We included two types of backfills: neutral soil and crushed limestone (here referred to as the backfill) to investigate if these materials were discriminated using ERT and IP data. Then, we defined two bodies of lime at different positions (x, z) and with different thicknesses to test the effect that including spatial trends has on the classification approaches. Finally, the upper limit of the bedrock has a step-like shape, which might be close to a limestone quarry structure. The resistivity and chargeability values that we used for the different regions are shown in Table 1. Then, similar to the field measurements, we created a dipole-dipole acquisition scheme with 64 electrodes spaced 1.5 m. For the numerical modeling of ERT and time-domain IP data sets, we used the open-source library pyBERT, which is based on the framework of pyGIMLI (Rücker et al., 2017).

The ERT data were modeled by adding a 3% voltage dependent noise plus 1 μV absolute error (e.g., Costall et al., 2020). Then, the

apparent chargeability was modeled following Seigel’s formulation, carrying out two DC resistivity forward models: the inverted resistivity of the medium and a decreased resistivity modified by the intrinsic chargeability (chargeability model, Table 1) (Seigel, 1959; Oldenburg and Li, 1994).

Finally, the synthetic data were inverted using the commercial software RES2DINV (Loke, 1997, 2004) to avoid the pitfall of using the same forward solver in the reconstruction algorithm (Lionheart, 2004). Here, we incorporated the data noise estimate for the apparent resistivity by subtracting the synthetic data modeled with and without added noise. We used a robust least-squares inversion with the Gauss-Newton method and an initial damping factor of 0.25. The inverted ERT and TDIP models are shown in Figure 2 together with the normalized sensitivity represented in the logarithmic scale and the real interfaces from Figure 1.

In the resistivity model, low values delineate the shallower lime deposit but the deeper lime cannot be imaged. There is no clear contrast of resistivity between the heterogeneous waste and the soil and backfill underneath. The upper limit of the bedrock is better imaged from $x \sim 20$ m to the end of the profile. Nevertheless, the resistivity values of the bedrock have a strong lateral variation as an effect of the regularization (there was a tradeoff between the damping factor and the root-mean-square [rms] error) and low sensitivity.

In the chargeability model, the surficial layer of waste is well delineated with large values. Nevertheless, the horizontal interface of the backfill/soil and the lime cannot be clearly distinguished. Artifacts of large chargeability are present at the locations of the lime and larger artifacts in the bedrock area centered at $x \sim 10$ and 30 m.

For the assessment of the inverted models’ reliability, we used the normalized sensitivity in logarithmic scale (hereafter referred to as the sensitivity), which shows how the data are influenced by the respective resistivity of the model cells. In Figure 2, we can observe a general sensitivity decrease with depth, particularly below the shallowest lime deposit.

In addition, we present the crossplots of the inverted resistivity (ρ) and chargeability I values (Figure 3). For comparison, we also plot the mean of the inverted data $\mu_i = (\rho_i, C_i)$ together with the initial values for modeling (Table 1) for each category. First, we can notice that the mean of the inverted data for the bedrock presents a largely underestimated value of resistivity compared with the initial value. The second category where we can see a large variation is the lime, where the mean of the inverted resistivity was larger than the initial modeling values. In addition, we notice that all the clusters’ categories are largely overlapping, especially the bedrock and lime, whose ρ and C values are widely distributed. This gives an insight

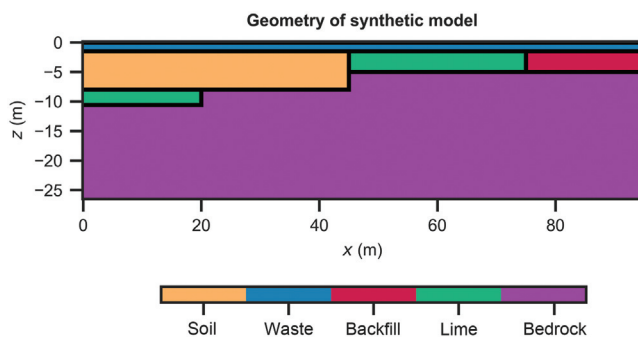


Figure 1. Geometry for the ERT and time-domain IP modeling.

Table 1. Resistivity and chargeability values for the different regions used for modeling.

Region marker	Material	Resistivity (Ωm)	Chargeability (mV/V)
1	Soil	300 (Wang et al., 2017)	15 (Kiberu, 2002)
2	Limestone	9000 (Sun et al., 2017)	10 (Johansson et al., 2017)
3	Backfill	800 (crushed limestone backfill, e.g., Qiao et al., 2019)	15 (North-West Europe, 2020)
4	Heterogeneous waste (nonorganic)	500 (Dumont et al., 2017)	100 (Elis et al., 2016)
5	Lime	3 (sample measured in the laboratory, this study)	1.3 (see Moreira et al. [2019] for chargeability values in dolomitic deposits, used as raw material for lime production)

into the ability of ERT and IP for resolving the features of this complex anthropogenic scenario and the uncertainty associated with the inversion process.

Synthetic borehole sampling

In this synthetic case, we assume a sampling scenario composed of several equidistant boreholes along the inverted 2D section.

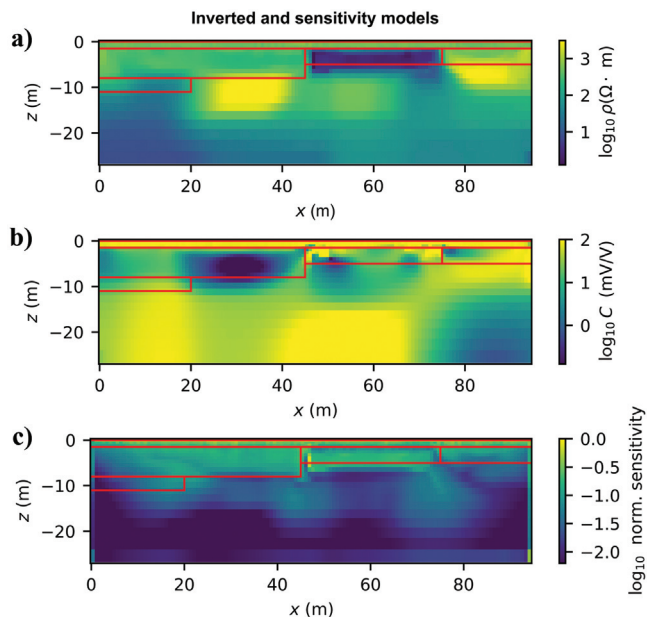


Figure 2. (a) Inverted resistivity and (b) chargeability models with the (c) associated normalized sensitivity. The rms errors of the inversion were 4.57% and 6.72%, respectively. Red boundaries represent the real model from Figure 1.

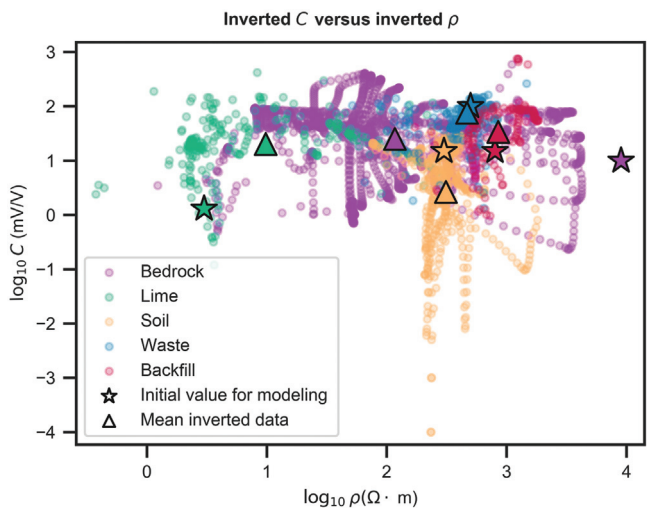


Figure 3. Crossplot of the inverted chargeability versus the inverted resistivity in logarithmic scale. The stars represent the original values for modeling, and the triangles represent the mean of the inverted data for each category, i.e., $\mu_i = (\rho_\mu, C_\mu)$. The inverted data corresponding to the different categories are represented with different colors.

Notice that in real study cases, we want to reduce the number of excavations to mitigate costs and environmental and health risks. Then, to select and justify an optimum number of boreholes we follow a statistical analysis. We used the mean of the inverted data for each class i , i.e., $\mu_i = (\rho_\mu, C_\mu)$ (see Figure 3, represented with triangles), and computed the mean of the inverted data within a variable number of boreholes b . For each class i , this was represented as $\mu_i(b)$. Then, we compute a summation of the difference between μ_i and $\mu_i(b)$ over all the classes i , i.e., $\sum_i(\mu_i - \mu_i(b))$. Figure 4 shows the plot of the summation versus the number of boreholes b . As b increases (and tends to cover the entire domain), the summation is closer to zero as $\mu_i(b) \rightarrow \mu_i$. The first points of this plot vary depending on the location of the selected boreholes, and for $b < 6$, the summation's rate decreases more significantly. Therefore, in the following sections, we start by using five boreholes equidistantly distributed (Figure 5) and test the effect of changing b in the probabilistic approach and the MLP model.

Interpretation — Classes prediction using a probabilistic approach

The training data were composed of $\{\rho, C, x, z\}$ at the borehole locations, which resulted in a matrix of 158×4 , and their respective known categories (vector of 158×1). In this case, we used a sensitivity threshold of $10^{-1.7}$ to keep only parts of the tomograms that are sufficiently reliable. The threshold was chosen based on the maxima of the sensitivity gradient, taking the average sensitivity values located beneath the area of the shallowest local maxima (below the conductive zone corresponding to the lime deposit). This threshold leads to the use of 38% of the data, which supports the assumption of testing this approach on heterogeneous models (few reliable data).

For each category observed in the boreholes' logs, $A_i = \text{soil, waste, backfill, lime, and bedrock}$, we estimate a prior probability $P(A_i)$ based on the area that they occupy on the boreholes (see Figure 5). The proportions of A_i are 9.57%, 5.5%, 2.5%, 7.3%, and 74.9%, respectively. Note that if the boreholes are limited in depth, one would naturally assume the vertical continuity of the bedrock once encountered. Then, we computed unimodal Gaussian

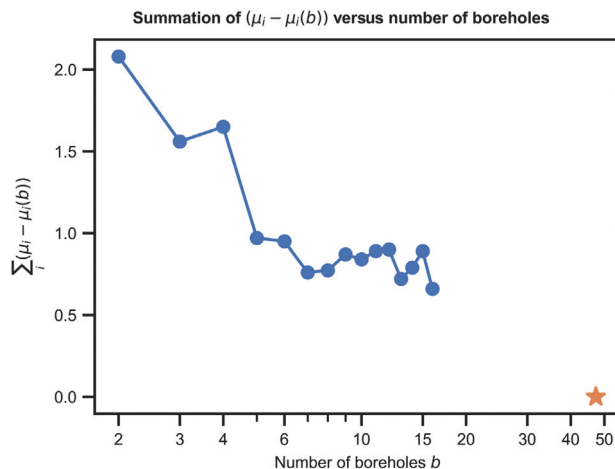


Figure 4. Summation of the difference between μ_i and $\mu_i(b)$ over all the classes versus the number of boreholes. Notice that initially $b = 2$ as we need to have all the category types.

distributions of the training data given each category A_i and computed the corresponding conditional probabilities according to equation 7. We selected this type of distribution as it can integrate the overall uncertainty from data noise, inversion artifacts, and scarcity of the collocated data. If a lot of collocated data are available, empirical distribution can be built without requiring any assumption about their shape. Figure 6 shows the conditional probabilities with and without considering the sensitivity threshold.

The distributions of Figure 6 show the impact of the large prior probability considered for the bedrock. This category presents the largest probability values for $\log_{10} \rho > 0.5$, for $\log_{10} C > 0$, at the largest depths $z < -6$ and nearly along the whole model domain in x . Yet, it can be observed that the resistivity model was able to discriminate between the bedrock and the lime. With the chargeability model, we can only discriminate the soil from the bedrock, due to the heterogeneities of large values distributed in the bedrock. This is the reason why it was not possible to clearly distinguish the waste (material with the largest chargeability modeling values) from the other categories.

The conditional probabilities given the spatial coordinates show a trend on the vertical (z) and lateral (x) distribution of the materials according to the location of the boreholes, i.e., the probabilities are impacted by the spatial distribution and the number of boreholes. Given the depths, several categories are clearly resolved: bedrock at the largest depths, the soil at intermediate depths, and waste at the shallowest zone. However, these distributions cannot discriminate between different categories at similar depths. In the distributions given the distance x , we can roughly differentiate between the soil (maximum probability at shorter x) and the backfill with larger probabilities at the end of the profile. Note the impact that the small sampled region of backfill has on its conditional probability (large probabilities in a reduced range of x).

Afterward, we computed the joint conditional probabilities for each category $P(A_i|\rho, C, x, z)$ using equation 4 and normalized them so that they sum to one. Hereafter, we refer to the normalized joint conditional probability as a joint probability. We work with the ratios of each data event, $g_i = (1 - P(A|G_i))/P(A|G_i)$, whose corresponding conditional probability can be independently estimated, i.e., $P(A_i|\rho)$, $P(A_i|C)$, $P(A_i|x)$, and $P(A_i|z)$. The results are presented as probability maps, see, for instance, Figure 7, where the joint probabilities of the waste and the bedrock are represented in the whole model domain. The surficial layer of waste was accurately delineated, whereas the upper limit of the bedrock was overestimated at the beginning of the profile $x < 20$ m.

We derived a map in terms of the categories by comparing the joint probabilities of the materials and selecting the category corresponding to the largest probability value (Figure 8). In this map, we added transparency which indicates the decrease in probability values: total opacity represents a probability of 1 and the strongest transparency represents a minimum probability of 0.25 for

four categories. Finally, we computed the classification scores comparing the results with the test data, i.e., categories of the synthetic model (excluding the data at the boreholes). We obtained an accuracy score of 0.87 and the confusion matrix is shown in Figure 8.

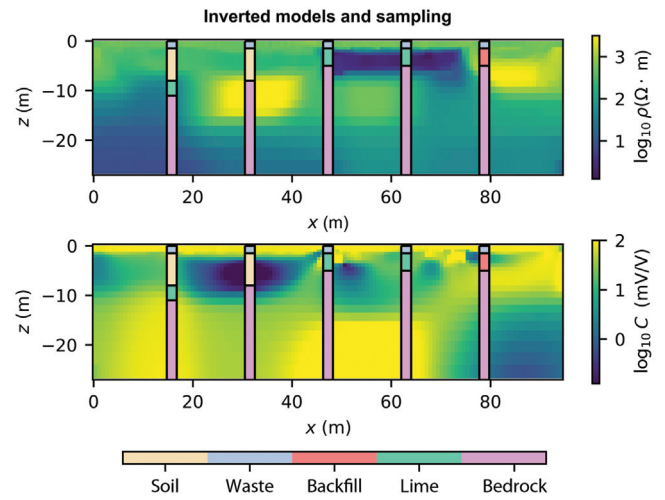


Figure 5. Resistivity and chargeability inverted models plotted with the boreholes. Where bedrock is present, we assume it reached the bottom of the models.

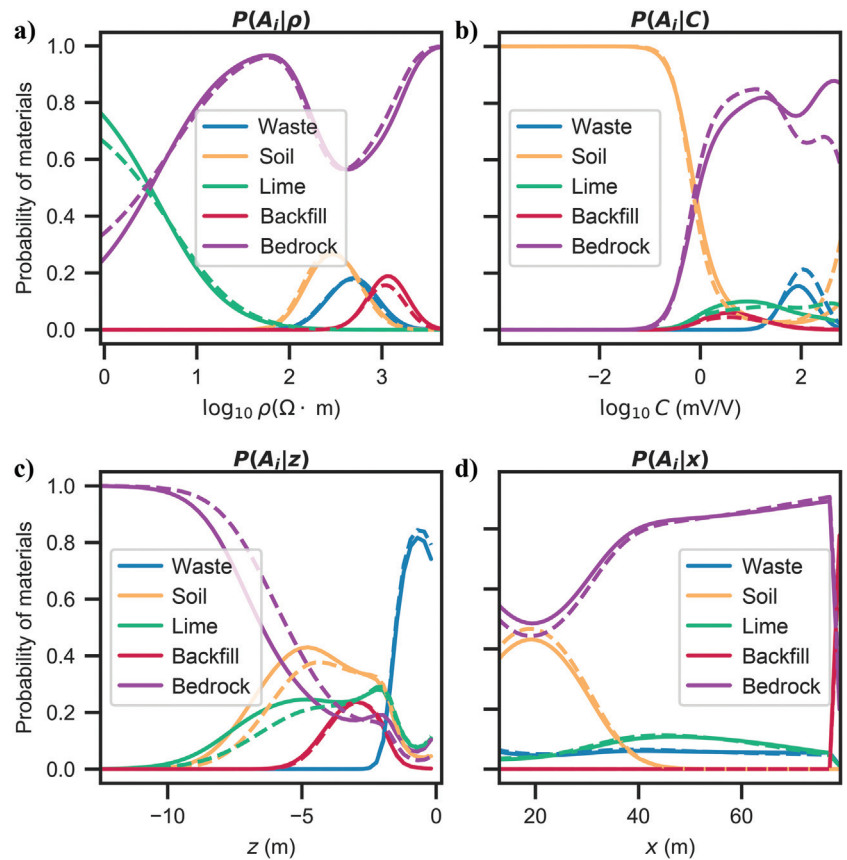


Figure 6. The conditional probability of A_i given (a and b) ρ and C , respectively, and (c and d) z and x from the boreholes. The solid lines represent the probabilities derived using the whole data within the boreholes and the dashed lines represent the distributions using a sensitivity threshold.

First, we can notice that the bedrock could be predicted along the entire model and especially in the deeper areas, although its occurrence was slightly overestimated. The waste deposit could be accurately delineated, whereas the soil and the lime deposit in the central part of the profile were roughly delimited laterally and vertically. This is represented in the confusion matrix, as we found that the bedrock obtained the largest number of correct predictions followed by the categories of waste, soil, and lime. The backfill could only be predicted in the area immediately around the borehole. This is

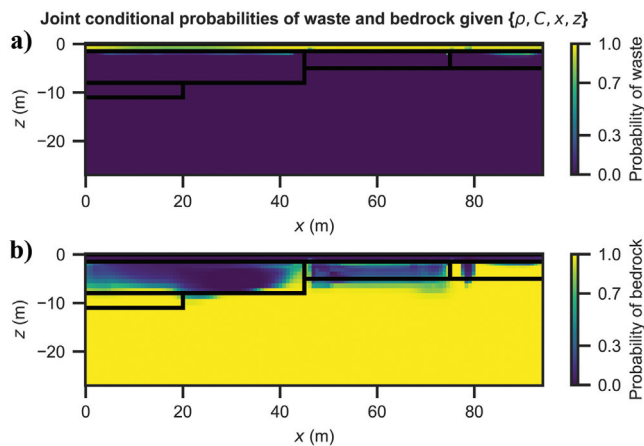


Figure 7. Joint probabilities of (a) the waste $P(A_{\text{waste}}|\rho, C, x, z)$ and (b) the bedrock, using a sensitivity threshold of -1.7 .

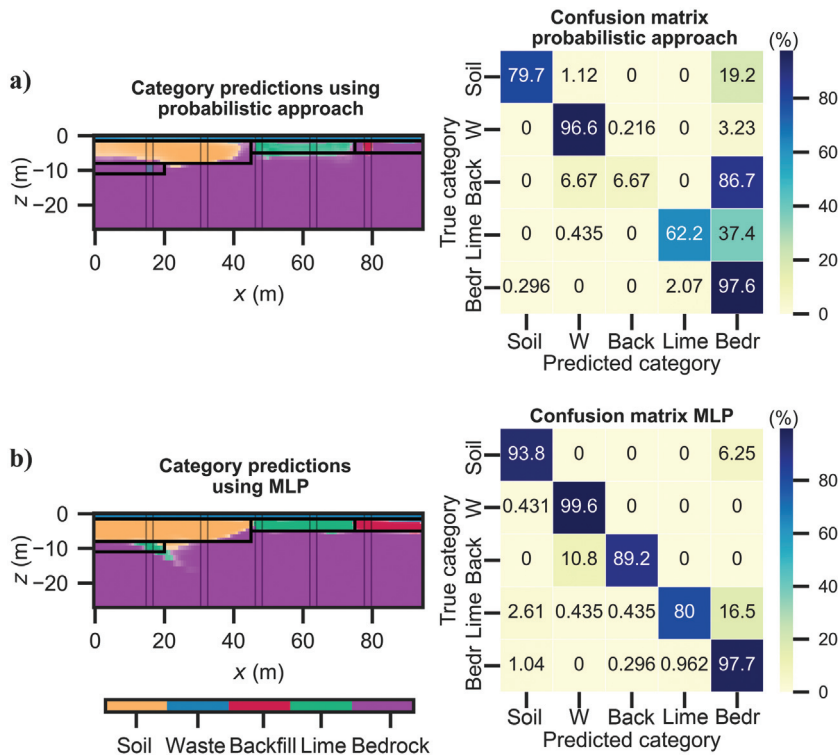


Figure 8. Category predictions from the (a) probabilistic approach and (b) MLP with borehole locations and corresponding confusion matrices. Minimum probability values (higher transparencies) are 38% for the probabilistic approach and 43% for MLP. In the confusion matrices, “w” refers to waste, “bedr” refers to bedrock, and “back” refers to backfill.

mostly an effect of integrating the horizontal tendency x of the boreholes in the joint probabilities, and this is the reason why in the confusion matrix the lowest percentage of correct classifications corresponds to backfill. The largest percentage of incorrectly predicted categories corresponded to the backfill which was misclassified as bedrock.

The second lime body, close to the origin of the profile, could not be detected, first because the inverted geophysical models could not resolve this feature and second because this category has considerably lower prior probability values compared with the bedrock. This is the reason why this zone was predicted as bedrock.

The transparency added in Figure 8 allows us to identify the zones where there is a larger uncertainty in defining the categories, e.g., the lateral interface between the soil and the lime.

Interpretation — Classes prediction using MLP

We apply the MLP algorithm using the python library of scikit-learn (Pedregosa et al., 2011). The training data were composed of $\{\rho, C, x, z\}$ at the borehole locations, resulting in a matrix of 158×4 , and their respective known categories (vector of 158×1). Similar to the probabilistic approach, we consider a sensitivity threshold of $10^{-1.7}$ on the selected training data. To optimize the architecture of MLP, we use a validation data set that is composed of 15% of the total training data and the remaining 85% of the data are used to train the algorithm. Several combinations of hyperparameters showed equally large accuracy scores. From the hyperparameters that showed the highest scores, we selected the simplest, two hidden layers of 100 neurons each, a regularization parameter of 0.15, a stochastic gradient-based optimizer, and the rectified linear unit as the activation function of the hidden layers.

At this step, we used the training data and the validation data to train the algorithm, i.e., the same input data as for the probabilistic approach. The category predictions in the whole model domain are shown in Figure 8, in which we also added transparency which represents the probability values from the activation function of the output layer (Softmax). Total opacity represents a probability of one, whereas total transparency represents a probability of zero. In Figure 8, we can see that the waste, the soil, one lime deposit, and the backfill were well delineated. The lime body located at the origin of the profile was partially imaged. To assess the performance of this algorithm, we also used the real model defined in Figure 1 as the test data set (excluding the data on the boreholes) and obtained an accuracy score of 0.95. Figure 8 shows the confusion matrix, where the largest numbers of correct predictions is the ones from waste, bedrock, and soil. The category that was incorrectly classified the most was the lime (predicted as bedrock). Yet, we can observe that the classification of the lime and backfill improved using MLP.

Effect of data sources as input

In this section, we show the effect the data sources or elements of the training data have

on the probabilistic approach and MLP. Figures 9, 10, and 11 show the category predictions of both methods and the respective confusion matrix when we use only the resistivity inverted data ρ colocated with the boreholes, the resistivity together with the chargeability inverted data ρ and C , and these two inverted data sets together with the depth z , respectively.

First, we can notice that when we only use ρ or ρ and C (with a sensitivity threshold of $10^{-1.7}$) as training data in MLP, the predictions are strongly influenced by the artifacts from the model inversion leading to several misclassified zones (see Figures 9 and 10). Some categories remain correctly identified such as the waste deposit and part of the soil. However, the probabilistic approach overestimates the distribution of the bedrock (largest prior probability). It partially distinguishes one lime deposit when using only ρ and delineates most of the waste layer and partially the soil when using ρ and C (Figures 9 and 10). Although the probabilistic approach is not able to predict all the categories in the whole domain, the results are still in line with the material proportions estimated from sampling and thus more realistic. This is not the case for MLP, where the waste, soil, backfill, and lime are predicted at larger depths.

When we use additionally the depth from the colocated boreholes (i.e., ρ , C , and z), the results of both methods largely improve (see Figure 11), and the categories are well delineated overall. Yet, the probabilistic approach indicates a larger uncertainty (more transparency) in the deposits of backfill, and partially, the soil. In addition, the classifications of both methods present few locations where the soil and backfill are misclassified. This is the improvement that we can observe when we use all the data sets: ρ , C , z , and x , especially for MLP where the soil is only predicted at $x < 50$ m and the backfill only for $x > 75$ m (see Figure 8).

The probabilistic approach presents some changes when using $\{\rho, C, z\}$ and $\{\rho, C, z, x\}$. The method proves better for classifying the backfill deposit when using only $\{\rho, C, z\}$ (accuracy score of 0.89). However, when x is included, the probabilistic approach is strongly impacted and, even though it reduces misclassification between the soil and backfill at a few locations, the backfill is only resolved roughly in the area of a borehole and the accuracy score is reduced to 0.87.

In general, including spatial information on the training data should be done carefully. In the probabilistic approach, the conditional probabilities are clearly impacted by the spatial distribution of the boreholes. Therefore, highly localized sampling of certain materials may lead to small classification zones in the immediate vicinity of a borehole. Although MLP presents large accuracy scores using the spatial training data, which indicates that the distribution of the boreholes reflected the real material distribution of the synthetic case, this is rarely the case in the field (Gahegan, 2000; Cracknell and Reading, 2014; Baasch et al., 2018), especially in heterogeneous landfills where abrupt vertical and lateral variations of materials may be present.

Effect of borehole sampling

In this section, we assess the effect that the number of available boreholes and their distribution has on both methods. We use the training data set $\{\rho, C, z\}$, which led to a larger accuracy score in the probabilistic approach as compared with $\{\rho, C, z, x\}$. First, we assume a sampling survey with boreholes uniformly distributed along the whole domain (Figure 5). Second, we assume a survey composed of boreholes whose distribution does not map all the model domains in the x - and z -directions, i.e., boreholes might be concentrated in an area of the model and/or boreholes might not go deep enough to map the bedrock upper interface (see Figure 12).

Figure 13 shows the plot of the accuracy score against the number of boreholes for both sampling scenarios. Despite the fact that the minimum number of boreholes to capture all categories is two, this number leads to a highly variable accuracy depending on the distribution of the boreholes. Yet, this variability is largely reduced when $b > 3$. In Figure 13, we can note that the accuracy scores of the probabilistic approach using the uniform or nonuniform sampling scenarios are very similar. However, MLP predicts correctly the classes at most locations of the model under the uniform borehole distribution. Nevertheless, if the boreholes do not cover the whole domain in x and z and, therefore, the training samples do not reflect the real distribution of the materials (preferential sampling), the classification performance decreases. In addition, we

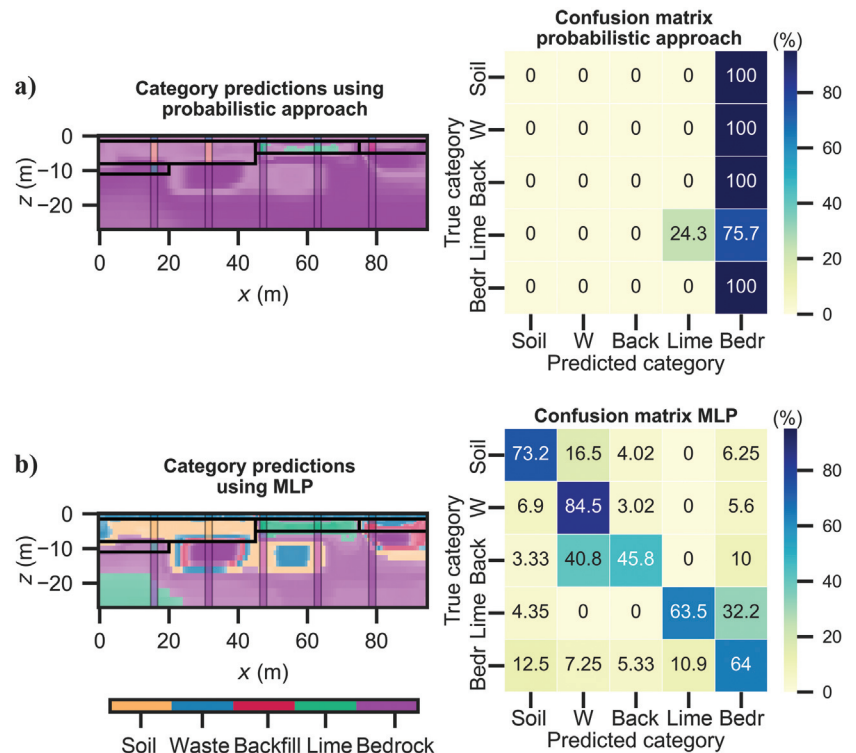


Figure 9. Category predictions of (a) probabilistic approach and (b) MLP algorithm using only ρ . The minimum probability values are 50% for the probabilistic approach and 40% for MLP, and the accuracy scores were 0.53 and 0.68, respectively. The confusion matrices also are displayed next to each model. The MLP model was built using two hidden layers with 100 neurons each, a regularization parameter of one, the rectified linear unit as an activation function of the hidden layers, and an optimizer based on the quasi-Newton method. In the confusion matrices, “w” refers to waste, “bedr” refers to bedrock, and back refers to backfill.

observed that MLP can lead to unrealistic classifications that underestimate the bedrock distribution even if the accuracy scores are similar to those of the probabilistic approach. See, for example, Figure 14, where we show the comparison between both approaches using three boreholes of a nonuniform sampling scenario (shown in Figure 12). Note that the lime deposit was predicted at the deepest regions of the model.

Field case study: Onoz landfill

Site description

The study site is in a former limestone quarry in Onoz (Walloon Region, Belgium) that produced lime until 1967. At the end of the quarry activities, the eastern part of the site was filled with slaked lime and fly ash. The area of interest here is the central part of the site, which was used as a landfill where different types of waste were deposited: inert waste, household, industrial waste, backfill, etc. (see Figure 15).

Since the landfill's closure, several sampling surveys mostly composed of trial pits and different geophysical measurements have been conducted. Here, we focus on one high-resolution 2D profile which has ERT and IP data (profile P2) and presents the largest number of collocated excavations where bedrock was reported. Profile P3 has only ERT data and profiles P1 and P3 present a very

similar distribution of resistivity and chargeability (Caterina et al., 2019).

Data acquisition and inversion

ERT and IP measurements were collected with an ABEM Terrameter LS. The profile presented here was acquired using 64 electrodes at 1.5 m electrode spacing. We used a dipole-dipole array configuration and a stack of $n = 2$ and a protocol sorted to limit electrode polarization, the electrical current was injected for 2 s, using an integration window of 1.7 s for the electrical resistance measurements, and the decay of electrical potential after current shut off was measured for 3 s.

Data were first filtered by removing the resistance measurements that presented variations larger than 5% from the measurements repeated two times. This is commonly referred to as the repetition error (e.g., Robert et al., 2011). On the IP data, the curves of inconsistent decay were also removed, which represented 18% of the original data points. Inversion of data was performed with RES2DINV (Loke, 1997, 2004). We also used a robust least-squares inversion with the Gauss-Newton method and an initial damping factor of 0.25. The inverted ERT and IP models are shown in Figure 16 together with the normalized relative sensitivity. We obtained an rms of 4.74% and 3.36% for the resistivity and chargeability models, respectively, after seven iterations.

The inverted resistivity section shows a conductive body on the top of a resistive horizon and a strong lateral contrast at approximately $x = 30$ m (Figure 16). The IP model presents scattered bodies of large chargeability on the surface and a smoother lateral contrast beneath. Similar to the synthetic model, the sensitivity of these inverted data presents a vertical decrease in the central part, below the conductive body.

Sampling

Several trial pits have been excavated in the landfill area (Figure 15). To test the probabilistic approach and the MLP algorithm, we consider the six trial pits that are collocated with the ERT/IP profile, and which do not cover the entire model domain. Only two shallow pits, not reaching the bedrock, are in the first half of the profile. We assume that once the bedrock is reached, it extends further at depth (in z) (see Figure 17).

Interpretation — Classes prediction using a probabilistic approach

As observed in the synthetic case, the training data set $\{\rho, C, z\}$ led to a larger accuracy score in the probabilistic approach and MLP. Therefore, here we interpreted the inverted models using as input data the values of $\{\rho, C, z\}$ collocated with the boreholes and ignore the variable x . This defines a matrix of 149×3 and a corresponding vector of categories of 149×1 . We divided the input data in a training data set (87%) and a test data set (13%) to assess the classification perfor-

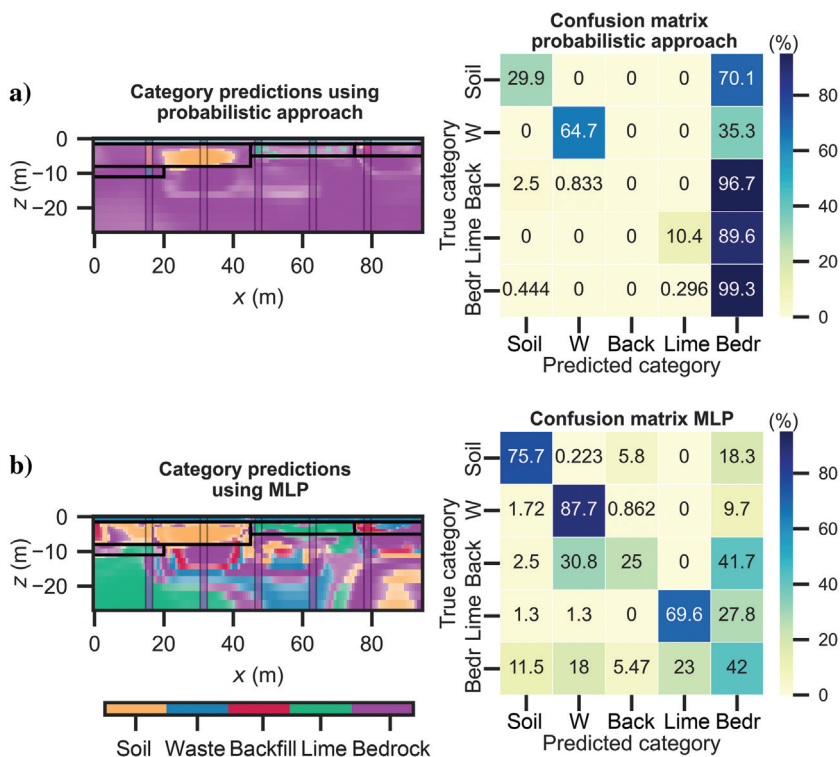


Figure 10. Category predictions of (a) probabilistic approach and (b) MLP algorithm using resistivity and chargeability data. For the former approach, minimum and maximum probability values are 44% and 100%, respectively, and 30% and 100% for MLP. The accuracy score for the probabilistic approach was 0.68 and 0.57 for MLP. The confusion matrices are displayed next to each model. The MLP model was built using three hidden layers with 50 neurons each, a regularization parameter of one, the rectified linear unit as an activation function of the hidden layers, and an optimizer based on the quasi-Newton method. In the confusion matrices, “w” refers to waste, “bedr” refers to bedrock, and back refers to backfill.

mance. First, we define the material proportions $P(A_i)$ based on the area of the trial pits. These were 11.33%, 1.78%, 4.67%, and 82.20% for waste, soil, lime, and bedrock, respectively. Then, we computed the joint probabilities $P(A_i|\rho, C, z)$ and derived the

classification maps. Note that the MLP algorithm intrinsically uses the prior distribution in its training step (if one class has a higher probability, the algorithm will more often classify an unknown point in that category).

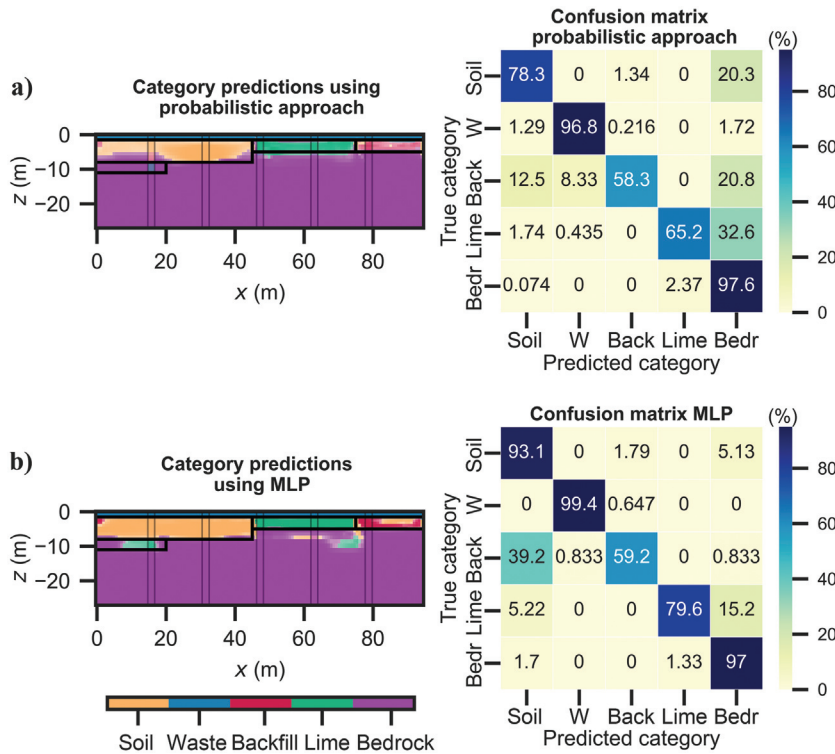


Figure 11. Category predictions of (a) probabilistic approach and (b) MLP algorithm using ρ , C , and the depth z . For the former, minimum and maximum probability values are 30% and 100%, respectively, and 41% and 100% for MLP. The accuracy score for the probabilistic approach was 0.89 and 0.93 for MLP. The confusion matrices also are displayed next to each model. The MLP model was built using three hidden layers with 100 neurons each, a regularization parameter of one, the rectified linear unit as an activation function of the hidden layers, and a stochastic gradient-based optimizer. In the confusion matrices, “w” refers to waste, “bedr” refers to bedrock, and back refers to backfill.

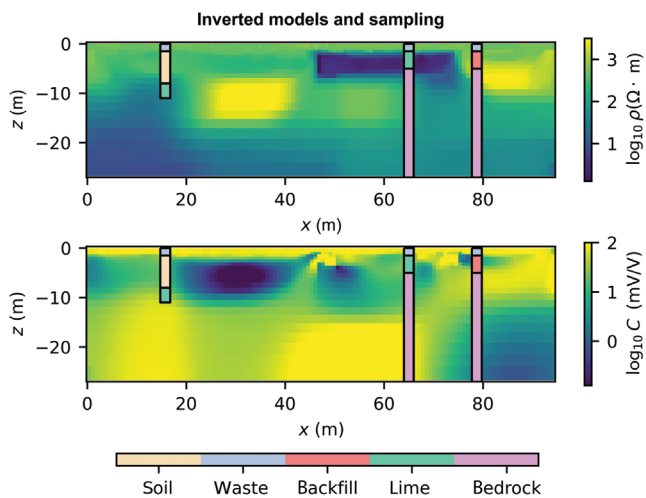


Figure 12. Nonuniform sampling scenario composed of three boreholes plotted with the inverted models.

Interpretation — Classes prediction using MLP

Similar to the probabilistic approach, the training data were the values of $\{\rho, C, z\}$ collocated with the boreholes, which defined a matrix of 149×3 and a corresponding vector of categories of 149×1 . The training data set was then divided into 72% of the total data, the validation data set to optimize the hyperparameters was 15%, and the test data set was the remaining 13%. From the three MLP architectures that showed the largest classification performance, we selected the simplest one. It was composed of one hidden layer with 50 neurons, a regularization value of 0.01, the rectified linear unit as an activation function of the hidden layers, and a quasi-Newton-based optimizer as the solver.

Comparison between the probabilistic approach and MLP

We use the same training data for the probabilistic approach and MLP (training and validation after tuning hyperparameters) and we use the same test data set to assess their performance. Figure 18 shows the category predictions for the probabilistic approach and MLP, the accuracy scores (which were 0.63 and 0.68, respectively), and the confusion matrices. In both approaches, the confusion matrix indicates that only the categories of waste and bedrock could be partially predicted.

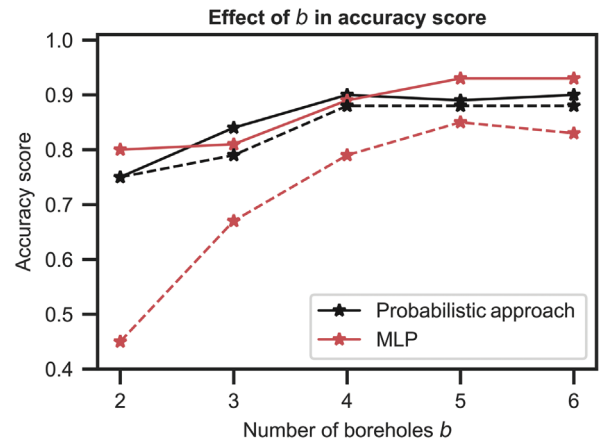


Figure 13. Plots of the accuracy score against the number of boreholes for the probabilistic approach (in black) and MLP (in red). The solid lines represent the sampling scenario of a uniform distribution of boreholes and the dashed lines represent the scenario of sparsely distributed boreholes.

Downloaded 05/25/23 to 134.157.32.210. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/page/policies/terms DOI:10.1190/geo2022-0133.1

First, we can observe that the MLP predicted model is strongly influenced by the inversion. The bedrock is predicted in the largest resistivity values and the lime is roughly predicted in the area of very low resistivities. The soil is predicted at some small areas of high chargeability (which might be artifacts in the inverted model), and as it was the class found at the largest depth of a pit at $x \sim 0$ m, then the soil was predicted in the area nearby at larger depths. This is a consequence of including spatial training data that are not distributed over the entire survey area (Baasch et al., 2018) and which may not reflect the real distribution of the materials deposited in the landfill. The waste also is predicted at larger depths in the model and might be influenced by the intermediate resistivity values. In addition, this model presents large areas of transparency representing probability values of approximately 30%. This means that there are large uncertainties in the prediction of classes in nearly the whole model except in the area of the predicted bedrock.

The classification derived from the probabilistic approach is composed of a much simpler model. It mainly presents three nearly continuous deposits: a surficial layer of waste (with interspersed soil), an underlying layer of lime, and the bedrock at the bottom, which is continuous along the whole model. Here, the zones of transparency are distributed in the shallowest layers along the profile, at a depth corresponding to the intersection between the lime and the waste. Nevertheless, this classification might present a more realistic geol-

ogy, with continuous bedrock and a nearly continuous surficial layer of heterogeneous waste on the top of a lime deposit. In general, the anthropogenic-geologic scenario from the probabilistic approach might be more realistic as it agrees with the additional trial pit logs excavated near the profile's origin, where bedrock was found at depths similar to those presented here.

DISCUSSION

We analyzed the performance of both approaches using a realistic synthetic benchmark and a field case where only few ground-truth data are available. First, the MLP algorithm requires a previous optimization of the hyperparameters using a validation data set (typically approximately 15% of the available data set). Thus, when little data are available, optimizing the hyperparameters with a validation set can be difficult and yield highly variable results. For instance, several MLP architectures for the synthetic case led to large accuracy scores using the validation set (which may be a consequence of having little data). Nonetheless, the use of the algorithm on new data requires again an optimization of hyperparameters using a validation data set (e.g., Aszemi and Dominic, 2019; Yu et al., 2020). Opositely, in the field case, the accuracy scores were, in general, smaller during the optimization of hyperparameters and only a few architectures presented scores of more than 0.7. In both cases, we selected relatively simple neural networks with a small number of hidden layers and neurons to remain in agreement with a small number of training data and to be comparable with the probabilistic approach where no data transformations are performed. In addition, during the optimization of hyperparameters, we observed larger accuracy scores in neural networks with smaller hidden layers and neurons.

Second, in the synthetic case, the entire model domain was available to assess the performance of the algorithm. This was not the case for the field site, where the test data were only 13% of the ground-truth data. For the latter, the classification predictions were compared with the observation from excavations nearby. Finally, the transparency applied in this classification is derived from the probabilities of the output layer and derived from the Softmax function. Note, however, that this function transforms the data into a probability vector with values between zero and one and whose elements add up to one. As it is a normalized exponential function, then the largest values are transformed into values close to one, whereas the smallest are into values close to zero. Therefore, the results may not represent exact probabilities as derived from the probabilistic approach (e.g., Gal and Ghahramani, 2016).

The probabilistic approach does not require an initial tuning of parameters. The prior probability of the different categories or the categories' proportion can be defined from the volumes obtained in the excavations. This information is essential to ensure that the conditional probabilities of each category are in line with the ground-truth data.

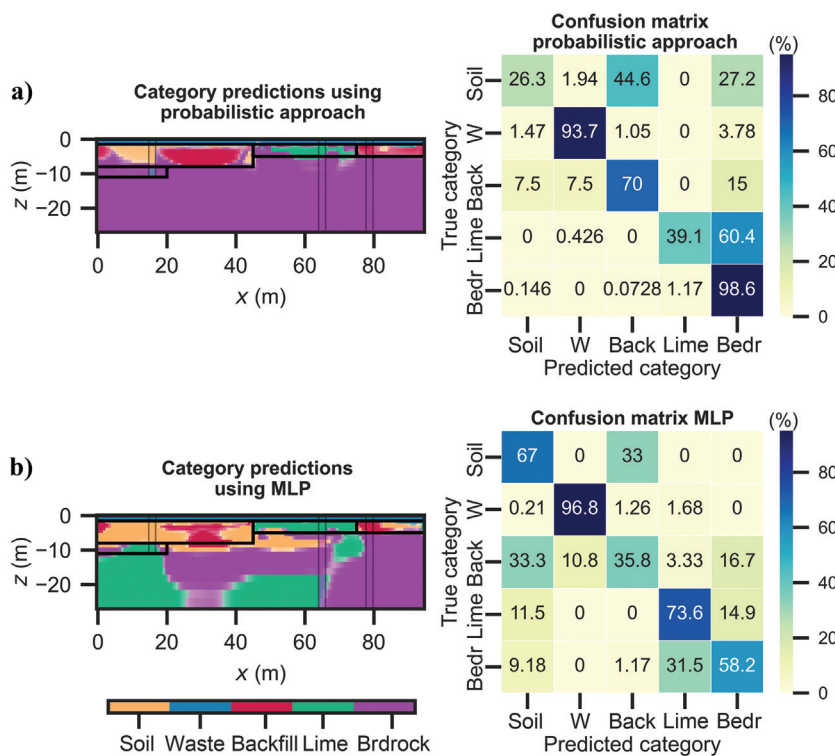


Figure 14. Category predictions of (a) probabilistic approach and (b) MLP algorithm using ρ , C , and z . The minimum probability values (higher transparencies) are 30% for the probabilistic approach, 45% for MLP, and 100% for the maximum probability values in both. The accuracy score for the probabilistic approach was 0.79 and 0.67 for MLP. The confusion matrices also are displayed next to each model. The MLP model was built using four hidden layers with 20 neurons, a regularization parameter of 0.1, the rectified linear unit as an activation function of the hidden layers, and a stochastic gradient-based optimizer. The distribution of nonuniform borehole sampling is shown in Figure 12. In the confusion matrices, “w” refers to waste, “bedr” refers to bedrock, and back refers to backfill.

If one category dominates the prior probability, whereas the geophysical data are poorly informative, this can lead to a final classification favoring the most probable facies. This is the reason why, in the field case, the bedrock lies across most of the model domain (especially in comparison with MLP classification). Finally, the transparency applied in this classification shows joint probabilities of the selected category, indicating specific zones of larger classification uncertainty. The probabilistic approach also is less dependent on the location and number of the collocated data.

The integration of more (inverted) data that increase the training data set is likely to improve the performance of MLP while not necessarily the performance of the probabilistic approach. When the physical properties of the categories in the inverted model(s) are highly variable along the whole domain mapped, the classification uncertainty is likely to increase. Nevertheless, the predictions would still be in line with the material proportions or prior information from sampling. Note that the better the materials are resolved in the conditional probabilities given the geophysical or spatial variables, the better the performance of the probabilistic approach. Consequently, in highly heterogeneous environments, the probabilistic approach is likely to improve when it is applied locally, i.e., per profile if multiple profiles are available or in areas of a 3D model, depending on the heterogeneity observed in the conditional probabilities. Regardless, a representative sampling based on geophysical data (e.g., Van De Vijver et al., 2019) could improve data

interpretation and therefore the classification performance of both methods.

Another option to increase the training data of the field case may be to use the data of the synthetic case. It is a common practice to create synthetic data to train artificial neural networks (Yu and Ma, 2021). Nevertheless, this contribution aims to present a probabilistic approach as an alternative to performing a rapid quantitative interpretation of site-specific anthropogenic environments incorporating ground-truth data.

We also studied the effect that the use of different data sources as training has on the probabilistic approach and MLP. Here, we can notice that the category predictions derived from MLP are highly influenced by the spatial heterogeneity of the inverted models when

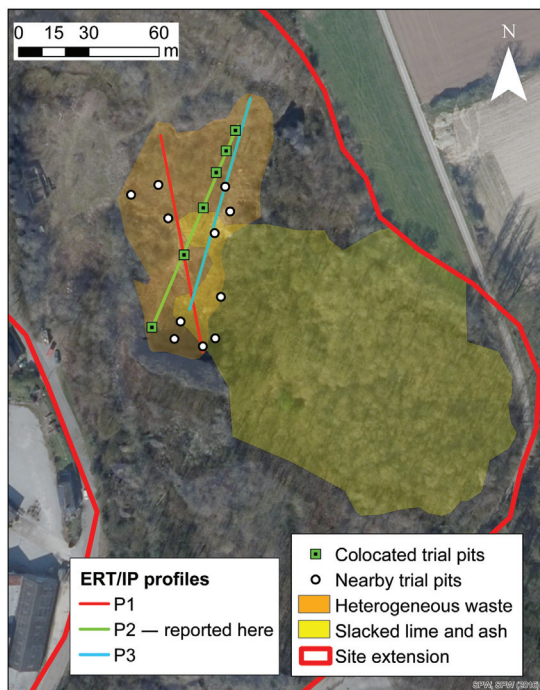


Figure 15. Map of the Onoz landfill, whose extension is indicated with the solid red line. The zone in orange is the lower part of the quarry, where heterogeneous waste was deposited, and the zone in yellow corresponds to the upper level of the quarry mainly composed of slacked lime and ash. P1, P2, and P3 show the locations of the ERT/IP profiles acquired in the lower part of the quarry. The square green symbols represent the position of trial pits collocated with a profile and used here, and the white circles are the trial pits excavated in the zone of the geophysical acquisition. The methodology that we present here was applied to profile P2 (in green) which is collocated with six trial pits.

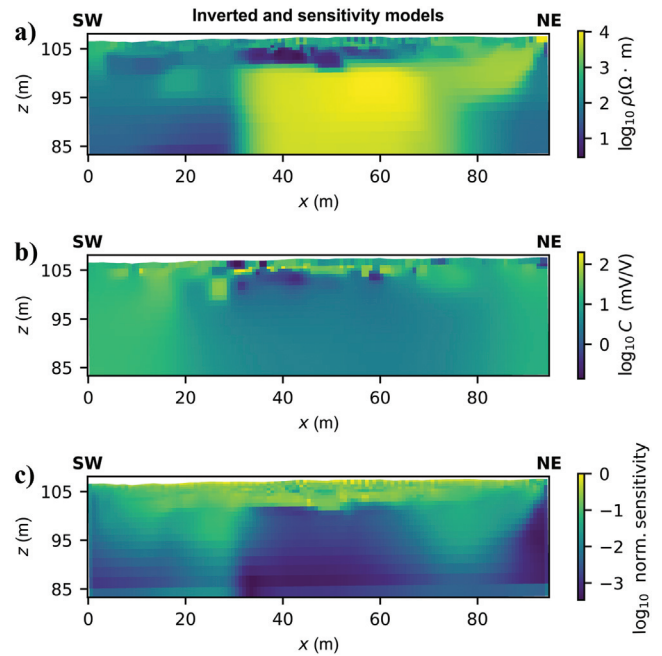


Figure 16. (a) Inverted resistivity model, (b) chargeability, and (c) associated sensitivity. The rms is 4.74% and 3.36% for the resistivity and chargeability models, respectively.

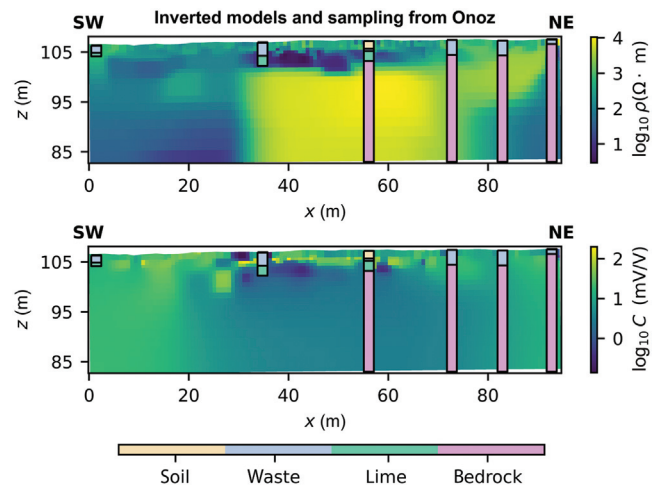


Figure 17. Collocated trial pits in the ERT and IP inverted sections.

we only use the geophysical data as input. This leads to several misclassification zones, most of which are nonrealistic. However, because the probabilistic approach relies on a Bayesian framework, it integrates the uncertainty related to data noise and inversion artifacts overall into the results. Therefore, the probabilistic approach is less sensitive to the heterogeneities of the inverted models when using only ρ and C , although it overestimates the distribution of the bedrock due to the large prior probability of this category. When we also include spatial training data (x, z), both approaches largely improve the classification, in particular MLP. However, including position x in the probabilistic approach leads to a high degree of influence on the sampled location.

In addition, we analyzed the effect that the number of boreholes and their distribution has on both approaches using the synthetic case. The distribution of a low number of boreholes can lead to variable accuracy scores (in particular for MLP). The accuracy scores of the probabilistic approach are very similar under the uniform and nonuniform sampling scenarios.

Here, we only compared the probabilistic approach with one algorithm of supervised learning, both of which presented (overall) similar classification models. Unsupervised learning also has been used for data interpretation in cases with minimal prior knowledge or where few ground-truth data are available (Delforge et al., 2021; Sabor et al., 2021; Whiteley et al., 2021). These approaches have proven useful for the interpretation of geophysical data in geologic environments composed of layered models or when the geophysical method(s) resolve zones or structures that can be evidenced in intrusive data. The motivation of this probabilistic approach is to quantitatively interpret

geophysical data in complex anthropogenic environments with extreme heterogeneity not only in terms of the spatial distribution of deposited wastes but also in terms of the high contrasts in physical properties that may lead to noisy data and artifacts in the inverted models. Because it provides probability values, its integration within other model types or in decision-making is relatively straightforward (e.g., Hermans et al., 2015).

The probabilistic approach and MLP classify zones with large uncertainty. If the classification improves when adding the x and z information, this also can lead to local improvement while degrading the overall accuracy score. Another option is then to improve the inverted model by adding the prior information from boreholes in the inversion process (Linde et al., 2015; Ronczka et al., 2015). For example, adding the depth of the bedrock, located in a low sensitivity zone, can improve the interfaces in the other part of the model (Caterina et al., 2014; Thibaut et al., 2021). However, adding such information bears the same limitation as it is highly dependent on the available boreholes and can lead to erroneous inverted models (Caterina et al., 2014). Nevertheless, using more advanced inversion methods does not prevent the use of the probabilistic approach or the MLP algorithm for post-inversion classification. Hermans and Irving (2017) show that an appropriate regularization could lead to an increase in the confidence of the classification (higher probabilities).

CONCLUSION

In this study, we presented a probabilistic approach that can be used in a classification problem including uncertainty estimation from site-specific multiple geophysical data sets and when only a few ground-truth data are available. The classification is based on two geophysical models (ERT and IP) and spatial data collocated with boreholes or trial pits. We compare this approach with a machine learning approach, the MLP, in a synthetic model and in a real field case. In addition, we tested the effects that the types of (training) data sources and the borehole sampling have on both approaches.

The probabilistic approach has proven to provide robust results regarding the position and number of ground-truth data and the presence of artifacts of inversion, in contrast to the MLP algorithm whose performance is largely related to the number of training data. Therefore, we recommend the use of the probabilistic approach in complex anthropogenic-geologic scenarios or other site-specific environments where: (1) geophysical inverted models present spatial heterogeneities (laterally and vertically) or artifacts, (2) only few ground-truth data are available, and (3) ground-truth data might not be sparsely distributed nor covering most of the area of study. The approach can be easily extended to integrate geophysical data from multiple methods or in three dimensions. It represents a suitable alternative to performing a rapid quantitative interpretation of geophysical data by using a probabilistic classification. Finally, as it integrates uncertainties in the prediction results, these can be used to complement decision support tools in sustainable landfill

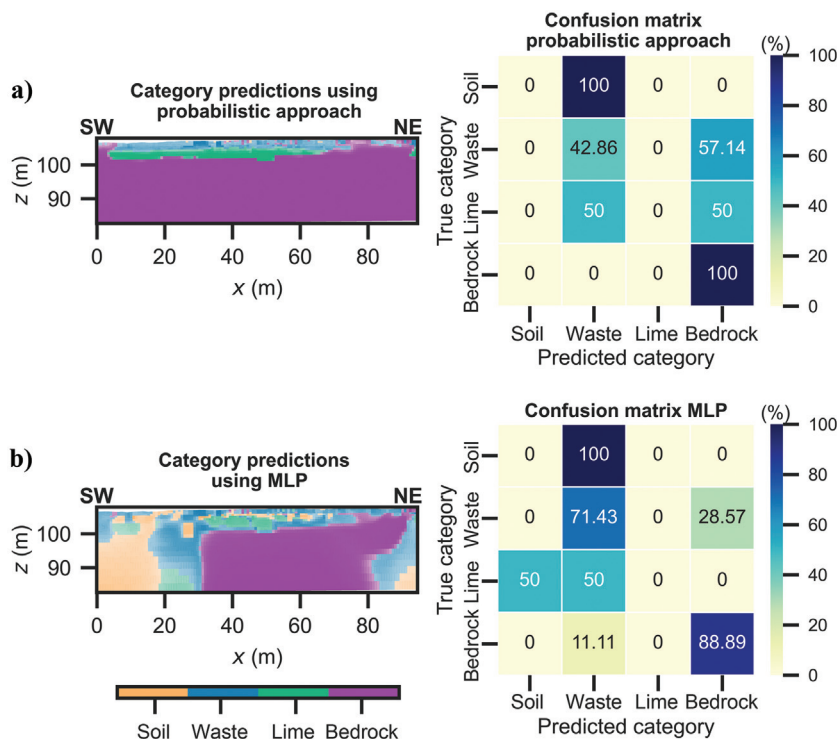


Figure 18. Category predictions derived for the (a) probabilistic approach and (b) MLP algorithm using ρ , C , and z . The minimum probability values (higher transparencies) are 35% for the probabilistic approach and 33% for MLP. The maximum probability value in both methods is 100%. Corresponding confusion matrices are shown next to category predictions.

management. In contrast, when a large number of training data are available, the MLP algorithm is expected to outperform the probabilistic approach.

ACKNOWLEDGMENTS

This work has been financed by the European Union's program of Interreg North-West Europe and the Wallon Region within the framework of the multidisciplinary projects of RAWFILL and NWE-REGENERATIS. We thank J. Whiteley, B. Baasch, an anonymous reviewer, and editors for their valuable comments and suggestions.

DATA AND MATERIALS AVAILABILITY

Codes necessary to reproduce the classification using the probabilistic approach in the synthetic case and in the field case study are available at <https://doi.org/10.5281/zenodo.7121021>.

REFERENCES

- Aster, R. C., B. Borchers, and C. H. Thurber, 2018, Parameter estimation and inverse problems: Elsevier.
- Aszemi, N. M., and P. D. D. Dominic, 2019, Hyperparameter optimization in convolutional neural network using genetic algorithms: *International Journal of Advanced Computer Science and Applications*, **10**, 269–278, doi: [10.14569/IJACSA.2019.0100638](https://doi.org/10.14569/IJACSA.2019.0100638).
- Baasch, B., H. Müller, and T. von Dobeneck, 2018, Predictive modelling of grain-size distributions from marine electromagnetic profiling data using end-member analysis and a radial basis function network: *Geophysical Journal International*, **214**, 460–473, doi: [10.1093/gji/ggy152](https://doi.org/10.1093/gji/ggy152).
- Beaujean, J., F. Nguyen, A. Kemna, A. Antonsson, and P. Engesgaard, 2014, Calibration of seawater intrusion models: Inverse parameter estimation using surface electrical resistivity tomography and borehole data: *Water Resources Research*, **50**, 6828–6849, doi: [10.1002/2013WR014020](https://doi.org/10.1002/2013WR014020).
- Bobé, C., E. Van De Vijver, J. Keller, D. Hanssens, M. Van Meirvenne, and P. De Smedt, 2020, Probabilistic 1-D inversion of frequency-domain electromagnetic data using a Kalman ensemble generator: *IEEE Transactions on Geoscience and Remote Sensing*, **58**, 3287–3297, doi: [10.1109/TGRS.2019.2953004](https://doi.org/10.1109/TGRS.2019.2953004).
- Caterina, D., T. Hermans, and F. Nguyen, 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](https://doi.org/10.3997/1873-0604.2013070).
- Caterina, D., I. Isonza Manrique, C. Inauen, A. Watlet, R. De Rijdt, G. Dumont, J. Chambers, and F. Nguyen, 2019, Contribution of geophysical methods to the study of old landfills: A case study in Onoz (Belgium): 17th International Waste Management and Landfill Symposium.
- Chasseriau, P., and M. Chouteau, 2003, 3D gravity inversion using a model of parameter covariance: *Journal of Applied Geophysics*, **52**, 59–74, doi: [10.1016/S0926-9851\(02\)00240-9](https://doi.org/10.1016/S0926-9851(02)00240-9).
- Costall, A. R., B. D. Harris, B. Teo, R. Schaa, F. M. Wagner, and J. P. Pigois, 2020, Groundwater throughflow and seawater intrusion in high quality coastal aquifers: *Scientific Reports*, **10**, 1–33, doi: [10.1038/s41598-020-66516-6](https://doi.org/10.1038/s41598-020-66516-6).
- Cracknell, M. J., and A. M. Reading, 2014, Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information: *Computers & Geosciences*, **63**, 22–33, doi: [10.1016/j.cageo.2013.10.008](https://doi.org/10.1016/j.cageo.2013.10.008).
- Delforge, D., A. Watlet, O. Kaufmann, M. Van Camp, and M. Vanclooster, 2021, Time-series clustering approaches for subsurface zonation and hydrofacies detection using a real time-lapse electrical resistivity dataset: *Journal of Applied Geophysics*, **184**, doi: [10.1016/j.jappgeo.2020.104203](https://doi.org/10.1016/j.jappgeo.2020.104203).
- Dewar, N., and R. Knight, 2020, Estimation of the top of the saturated zone from airborne electromagnetic data: *Geophysics*, **85**, no. 5, EN63–EN76, doi: [10.1190/geo2019-0539.1](https://doi.org/10.1190/geo2019-0539.1).
- Dramsch, J. S., 2020, 70 years of machine learning in geoscience in review: *Advances in Geophysics*, **61**, 1–55, doi: [10.1016/bs.agph.2020.08.002](https://doi.org/10.1016/bs.agph.2020.08.002).
- Dumont, G., T. Robert, N. Marck, and F. Nguyen, 2017, Assessment of multiple geophysical techniques for the characterization of municipal waste deposit sites: *Journal of Applied Geophysics*, **145**, 74–83, doi: [10.1016/j.jappgeo.2017.07.013](https://doi.org/10.1016/j.jappgeo.2017.07.013).
- Elis, V. R., A. T. Ustra, M. C. Hidalgo-Gato, O. J. Pejon, and F. Y. Hiodo, 2016, Application of induced polarization and resistivity to the environmental investigation of an old waste disposal area: *Environmental Earth Sciences*, **75**, 1–13, doi: [10.1007/s12665-016-6157-5](https://doi.org/10.1007/s12665-016-6157-5).
- Fossum, K., S. Alyaev, J. Tveranger, and A. H. Elsheikh, 2022, Verification of a real-time ensemble-based method for updating earth model based on GAN: *Journal of Computational Science*, **65**, 101876, doi: [10.1016/j.jocs.2022.101876](https://doi.org/10.1016/j.jocs.2022.101876).
- Fu, Q., D. Niu, Z. Zang, J. Huang, and L. Diao, 2019, Multi-stations' weather prediction based on hybrid model using 1D CNN and bi-LSTM: *Chinese Control Conference (CCC)*, 3771–3775, doi: [10.23919/ChiCC.2019.8866496](https://doi.org/10.23919/ChiCC.2019.8866496).
- Gahegan, M., 2000, On the application of inductive machine learning tools to geographical analysis: *Geographical Analysis*, **32**, doi: [10.1111/j.1538-4632.2000.tb00420.x](https://doi.org/10.1111/j.1538-4632.2000.tb00420.x).
- Gal, Y., and Z. Ghahramani, 2016, Dropout as a Bayesian approximation: Representing model uncertainty in deep learning: *International Conference on Machine Learning*, 1050–1059.
- Goodfellow, I., Y. Bengio, and A. Courville, 2016, *Deep learning*: MIT Press.
- Hellman, K., M. Roncza, T. Günther, M. Wennermark, C. Rücker, and T. Dahlin, 2017, Structurally coupled inversion of ERT and refraction seismic data combined with cluster-based model integration: *Journal of Applied Geophysics*, **143**, 169–181, doi: [10.1016/j.jappgeo.2017.06.008](https://doi.org/10.1016/j.jappgeo.2017.06.008).
- Hermans, T., and J. Irving, 2017, Facies discrimination with electrical resistivity tomography using a probabilistic methodology: Effect of sensitivity and regularization: *Near Surface Geophysics*, **15**, 13–25, doi: [10.3997/1873-0604.2016047](https://doi.org/10.3997/1873-0604.2016047).
- Hermans, T., F. Nguyen, and J. Caers, 2015, Uncertainty in training image-based inversion of hydraulic head data constrained to ERT data: *Workshop and case study: Water Resources Research*, **51**, 5332–5352, doi: [10.1002/2014WR016460](https://doi.org/10.1002/2014WR016460).
- Hermans, T., A. Vandenbohede, L. Lebbe, and F. Nguyen, 2012, A shallow geothermal experiment in a sandy aquifer monitored using electric resistivity tomography: *Geophysics*, **77**, no. 1, B11–B21, doi: [10.1190/geo2011-0199.1](https://doi.org/10.1190/geo2011-0199.1).
- Inauen, C. M., A. Brooks, D. Caterina, J. E. Chambers, B. Dashwood, A. Dimech, D. A. Gunn, I. Isonza Manrique, O. Neal, X. Piquet, D. Scott, A. Watlet, J. S. Whiteley, and P. B. Wilkinson, 2020, Combining an integrated geophysical survey into a landfill model: A case study from Emersons Green, UK: Presented at the 22nd EGU General Assembly.
- Johansson, S., C. Sparrenbom, G. Fiandaca, A. Lindskog, P. I. Olsson, T. Dahlin, and H. Rosqvist, 2017, Investigations of a Cretaceous limestone with spectral induced polarization and scanning electron microscopy: *Geophysical Journal International*, **208**, 954–972, doi: [10.1093/gji/ggw432](https://doi.org/10.1093/gji/ggw432).
- Jones, P. T., E. Wille, and J. Krook, 2018, 5 lessons learned, why we need to develop a broad dynamic landfill management strategy and vision for Europe's 500,000 landfills: 2nd EFM Seminar in the European Parliament, 1–12.
- Journé, A. G., 2002, Combining knowledge from diverse sources: An alternative to traditional data independence hypotheses: *Mathematical Geology*, **34**, 573–596, doi: [10.1023/A:1016047012594](https://doi.org/10.1023/A:1016047012594).
- Kaipio, J. P., V. Kolehmainen, M. Vauhkonen, and E. Somersalo, 1999, Inverse problems with structural prior information: *Inverse Problems*, **15**, 713–729, doi: [10.1088/0266-5611/15/3/306](https://doi.org/10.1088/0266-5611/15/3/306).
- Kiberu, J., 2002, Induced polarization and resistivity measurements on a suite of near surface soil samples and their empirical relationship to selected measured engineering parameters: M.S. thesis, ITC.
- Linde, N., P. Renard, T. Mukerji, and J. Caers, 2015, Geological realism in hydrogeological and geophysical inverse modeling: A review: *Advances in Water Resources*, **86**, 86–101, doi: [10.1016/j.advwatres.2015.09.019](https://doi.org/10.1016/j.advwatres.2015.09.019).
- Lionheart, W. R. B., 2004, EIT reconstruction algorithms: Pitfalls, challenges and recent developments: *Physiological Measurement*, **25**, 125–142, doi: [10.1088/0967-3334/25/1/021](https://doi.org/10.1088/0967-3334/25/1/021).
- Liu, L., Y. Liu, T. Li, Y. He, Y. Du, and Y. Luo, 2021, Inversion of vehicle-induced signals based on seismic interferometry and recurrent neural networks: *Geophysics*, **86**, no. 3, Q37–Q45, doi: [10.1190/geo2020-0498.1](https://doi.org/10.1190/geo2020-0498.1).
- Loke, M., 1997, *Res2DINV Software User's Manual*: University Sains Malaysia.
- Loke, M., 2004, *Tutorial: 2-D and 3-D electrical imaging surveys*: Geotomo Software.
- Lopez-Alvis, J., E. Laloy, F. Nguyen, and T. Hermans, 2021, Deep generative models in inversion: A review and development of a new approach based on a variational autoencoder: *Computers & Geosciences*, **152**, doi: [10.1016/j.cageo.2021.104762](https://doi.org/10.1016/j.cageo.2021.104762).
- Lopez-Alvis, J., F. Nguyen, M. C. Looms, and T. Hermans, 2022, Geophysical inversion using a variational autoencoder to model an assembled spatial prior uncertainty: *Journal of Geophysical Research: Solid Earth*, **127**, doi: [10.1029/2021JB022581](https://doi.org/10.1029/2021JB022581).
- Lysdahl, A. K., C. W. Christensen, A. A. Pfaffhuber, M. Vöge, L. Andresen, G. H. Skurdal, and M. Panzner, 2022, Integrated bedrock model combining airborne geophysics and sparse drillings based on an artificial neural network: *Engineering Geology*, **297**, doi: [10.1016/j.enggeo.2021.106484](https://doi.org/10.1016/j.enggeo.2021.106484).

- Magiera, T., B. Żogała, M. Szuszkiewicz, J. Pierwoła, and M. M. Szuszkiewicz, 2019, Combination of different geophysical techniques for the location of historical waste in the Izery Mountains (SW Poland): *Science of the Total Environment*, **682**, 226–238, doi: [10.1016/j.scitotenv.2019.05.180](https://doi.org/10.1016/j.scitotenv.2019.05.180).
- Moghadas, D., and A. Badorreck, 2019, Machine learning to estimate soil moisture from geophysical measurements of electrical conductivity: *Near Surface Geophysics*, **17**, 181–195, doi: [10.1002/nsg.12036](https://doi.org/10.1002/nsg.12036).
- Moreira, C. A., E. G. Dos Santos, L. M. Ilha, and R. Paes, 2019, Recognition of sulfides zones in Marble Mine through comparative analysis of electrical tomography arrangements: *Pure and Applied Geophysics*, **176**, 4907–4920, doi: [10.1007/s00024-019-02243-y](https://doi.org/10.1007/s00024-019-02243-y).
- North-West Europe, 2020, Co-funded by the Walloon region Landfill characterization: The landfill of Onoz (Wallonia, Belgium), https://www.nweurope.eu/media/10049/onoz_draft_website-version.pdf.
- Oldenburg, D. W., and Y. Li, 1994, Inversion of induced polarization data: *Geophysics*, **59**, 1327–1341, doi: [10.1190/1.1443692](https://doi.org/10.1190/1.1443692).
- Paasche, H., J. Tronicke, K. Holliger, A. G. Green, and H. Maurer, 2006, Integration of diverse physical-property models: Subsurface zonation and petrophysical parameter estimation based on fuzzy c-means cluster analyses: *Geophysics*, **71**, no. 3, H33–H44, doi: [10.1190/1.2192927](https://doi.org/10.1190/1.2192927).
- Parsekian, A. D., D. Grana, F. D. A. Neves, M. S. Pleasants, M. Seyfried, B. G. Moravec, J. Chorover, A. M. Moraes, N. Y. Smeltz, J. H. Westenhoff, and T. Kellners, 2021, Hydrogeophysical comparison of hillslope critical zone architecture for different geologic substrates: *Geophysics*, **86**, no. 5, WB29–WB49, doi: [10.1190/geo2020-0438.1](https://doi.org/10.1190/geo2020-0438.1).
- Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and É. Duchesnay, 2011, Scikit-learn: Machine learning in Python: *Journal of Machine Learning Research*, **12**, 2825–2830.
- Piana Agostinetti, N., and T. Bodin, 2018, Flexible coupling in joint inversions: A Bayesian structure decoupling algorithm: *Journal of Geophysical Research: Solid Earth*, **123**, 8798–8826, doi: [10.1029/2018JB016079](https://doi.org/10.1029/2018JB016079).
- Potdar, K., T. Pardawala, and C. D. Pai, 2017, A comparative study of categorical variable encoding techniques for neural network classifiers: *International Journal of Computer Applications*, **175**, 7–9, doi: [10.5120/ijca2017915495](https://doi.org/10.5120/ijca2017915495).
- Qiao, C., M. K. Moradillo, H. Hall, M. T. Ley, and W. J. Weiss, 2019, Electrical resistivity and formation factor of air-entrained concrete: *ACI Materials Journal*, **116**, 85–93, doi: [10.14359/51714506](https://doi.org/10.14359/51714506).
- Ray, A., S. Kaplan, J. Washbourne, and U. Albertin, 2018, Low frequency full waveform seismic inversion within a tree based Bayesian framework: *Geophysical Journal International*, **212**, 522–542, doi: [10.1093/gji/ggx428](https://doi.org/10.1093/gji/ggx428).
- Robert, T., A. Dassargues, S. Brouyère, O. Kaufmann, V. Hallet, and F. Nguyen, 2011, Assessing the contribution of electrical resistivity tomography (ERT) and self-potential (SP) methods for a water well drilling program in fractured/karstified limestones: *Journal of Applied Geophysics*, **75**, 42–53, doi: [10.1016/j.jappgeo.2011.06.008](https://doi.org/10.1016/j.jappgeo.2011.06.008).
- Romero-Ruiz, A., N. Linde, T. Keller, and D. Or, 2018, A review of geophysical methods for soil structure characterization: *Reviews of Geophysics*, **56**, 672–697, doi: [10.1029/2018RG000611](https://doi.org/10.1029/2018RG000611).
- Roncza, M., T. Voß, and T. Günther, 2015, Cost-efficient imaging and monitoring of saltwater in a shallow aquifer by using long electrode ERT: *Journal of Applied Geophysics*, **122**, 202–209, doi: [10.1016/j.jappgeo.2015.08.014](https://doi.org/10.1016/j.jappgeo.2015.08.014).
- Rücker, C., T. Günther, and F. M. Wagner, 2017, pyGIMLi: An open-source library for modelling and inversion in geophysics: *Computers & Geosciences*, **109**, 106–123, doi: [10.1016/j.cageo.2017.07.011](https://doi.org/10.1016/j.cageo.2017.07.011).
- Sabor, K., D. Jougnot, R. Guerin, B. Steck, J. M. Henault, L. Apffel, and D. Vautrin, 2021, A data mining approach for improved interpretation of ERT inverted sections using the DBSCAN clustering algorithm: *Geophysical Journal International*, **225**, 1304–1318, doi: [10.1093/gji/ggab023](https://doi.org/10.1093/gji/ggab023).
- Scheidt, C., L. Li, and J. Caers, 2018, Quantifying uncertainty in subsurface systems: John Wiley and Sons.
- Seigel, H. O., 1959, Mathematical formulation and type curves for induced polarization: *Geophysics*, **24**, 547–565, doi: [10.1190/1.1438625](https://doi.org/10.1190/1.1438625).
- Slater, L., and A. Binley, 2021, Advancing hydrological process understanding from long-term resistivity monitoring systems: *Wiley Interdisciplinary Reviews: Water*, **8**, e1513, doi: [10.1002/wat2.1513](https://doi.org/10.1002/wat2.1513).
- Sun, H., M. Cheng, C. Su, H. Li, G. Zhao, M. Su, S. Li, B. Zhang, L. Zhang, and K. Li, 2017, Characterization of shallow karst using electrical resistivity imaging in a limestone mining area: *Environmental Earth Sciences*, **76**, 1–9, doi: [10.1007/s12665-017-7112-9](https://doi.org/10.1007/s12665-017-7112-9).
- Thibaut, R., T. Kremer, A. Royen, B. Kim Ngun, F. Nguyen, and T. Hermans, 2021, A new workflow to incorporate prior information in minimum gradient support (MGS) inversion of electrical resistivity and induced polarization data: *Journal of Applied Geophysics*, **187**, doi: [10.1016/j.jappgeo.2021.104286](https://doi.org/10.1016/j.jappgeo.2021.104286).
- Van De Vijver, E., C. Bobe, D. Hanssens, P. De Smedt, and M. Van Meirvenne, 2019, Representative sampling of landfills: A robust procedure for selecting trench locations based on electromagnetic induction survey data: Presented at the AGU Fall Meeting 2019.
- Van De Vijver, E., D. Caterina, I. Isunza Manrique, C. Bobe, and F. Nguyen, 2020, Geophysical surveys for unlocking landfill resources: From past applications to future developments: Presented at the AGU Fall Meeting 2020.
- Van De Vijver, E., I. Isunza Manrique, C. Bobe, D. Caterina, T. Hermans, E. Wille, and F. Nguyen, 2021, Geophysics in support of dynamic landfill management: Moving beyond the challenges: Presented at the First International Meeting for Applied Geoscience & Energy, SEG, Expanded Abstracts, 3140–3144, doi: [10.1190/segam2021-3594435.1](https://doi.org/10.1190/segam2021-3594435.1).
- Wang, J., X. Zhang, and L. Du, 2017, A laboratory study of the correlation between the thermal conductivity and electrical resistivity of soil: *Journal of Applied Geophysics*, **145**, 12–16, doi: [10.1016/j.jappgeo.2017.07.009](https://doi.org/10.1016/j.jappgeo.2017.07.009).
- Wang, Z., H. Di, M. A. Shafiq, Y. Alaudah, and G. AlRegib, 2018, Successful leveraging of image processing and machine learning in seismic structural interpretation: A review: *The Leading Edge*, **37**, 451–461, doi: [10.1190/tle37060451.1](https://doi.org/10.1190/tle37060451.1).
- Wellmann, J. F., M. De la Varga, R. E. Murdie, K. Gessner, and M. Jessell, 2018, Uncertainty estimation for a geological model of the Sandstone greenstone belt, Western Australia — Insights from integrated geological and geophysical inversion in a Bayesian inference framework: Geological Society, London, Special Publications 453.
- Whiteley, J. S., J. E. Chambers, S. Uhlemann, P. B. Wilkinson, and J. M. Kendall, 2019, Geophysical monitoring of moisture-induced landslides: A review: *Reviews of Geophysics*, **57**, 106–145, doi: [10.1029/2018RG000603](https://doi.org/10.1029/2018RG000603).
- Whiteley, J. S., A. Watlet, S. Uhlemann, P. Wilkinson, J. P. Boyd, C. Jordan, J. M. Kendall, and J. E. Chambers, 2021, Rapid characterisation of landslide heterogeneity using unsupervised classification of electrical resistivity and seismic refraction surveys: *Engineering Geology*, **290**, doi: [10.1016/j.enggeo.2021.106189](https://doi.org/10.1016/j.enggeo.2021.106189).
- Williams, C. K. I., and D. Barber, 1998, Bayesian classification with Gaussian processes: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**, 1342–1351, doi: [10.1109/34.735807](https://doi.org/10.1109/34.735807).
- Yannah, M., K. Martens, M. Van Camp, and K. Walraevens, 2017, Geophysical exploration of an old dumpsite in the perspective of enhanced landfill mining in Kermt area, Belgium: *Bulletin of Engineering Geology and the Environment*, **78**, 55–67, doi: [10.1007/s10064-017-1169-2](https://doi.org/10.1007/s10064-017-1169-2).
- Yu, H., G. Chen, and H. Gu, 2020, A machine learning methodology for multivariate pore-pressure prediction: *Computers & Geosciences*, **143**, 104548, doi: [10.1016/j.cageo.2020.104548](https://doi.org/10.1016/j.cageo.2020.104548).
- Yu, S., and J. Ma, 2021, Deep learning for geophysics: Current and future trends: *Reviews of Geophysics*, **59**, doi: [10.1029/2021RG000742](https://doi.org/10.1029/2021RG000742).

Biographies and photographs of the authors are not available.