



HAL
open science

Convolutional neural networks with SegNet architecture applied to three-dimensional tomography of subsurface electrical resistivity: CNN-3D-ERT

M Vu, Abderrahim Jardani

► To cite this version:

M Vu, Abderrahim Jardani. Convolutional neural networks with SegNet architecture applied to three-dimensional tomography of subsurface electrical resistivity: CNN-3D-ERT. *Geophysical Journal International*, 2021, 225 (2), pp.1319-1331. 10.1093/gji/ggab024 . insu-03958483

HAL Id: insu-03958483

<https://insu.hal.science/insu-03958483>

Submitted on 16 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Convolutional neural networks with SegNet architecture applied to three-dimensional tomography of subsurface electrical resistivity: CNN-3D-ERT

M. T. Vu  and A. Jardani

Morphodynamique Continentale et Côtière, DeepGeoLearning consortium, Université de Rouen, M2C, UMR 6143, CNRS, 76821 Mont Saint Aignan, France.
E-mail: minh-tan.vu@univ-rouen.fr

Accepted 2021 January 19. Received 2021 January 15; in original form 2020 July 17

SUMMARY

In general, the inverse problem of electrical resistivity tomography (ERT) is treated using a deterministic algorithm to find a model of subsurface resistivity that can numerically match the apparent resistivity data acquired at the ground surface and has a smooth distribution that has been introduced as *prior* information. In this paper, we propose a new deep learning algorithm for processing the 3-D reconstruction of ERT. This approach relies on the approximation of the inverse operator considered as a nonlinear function linking the section of apparent resistivity as input and the underground distribution of electrical resistivity as output. This approximation is performed with a large amount of known data to obtain an accurate generalization of the inverse operator by identifying during the learning process a set of parameters assigned to the neural networks. To train the network, the subsurface resistivity models are theoretically generated by a geostatistical anisotropic Gaussian generator, and their corresponding apparent resistivity by solving numerically 3-D Poisson's equation. These data are formed in a way to have the same size and trained on the convolutional neural networks with SegNet architecture containing a three-level encoder and decoder network ending with a regression layer. The encoders including the convolutional, max-pooling and nonlinear activation operations are sequentially performed to extract the main features of input data in lower resolution maps. On the other side, the decoders are dedicated to upsampling operations in concatenating with feature maps transferred from encoders to compensate the loss of resolution. The tool has been successfully validated on different synthetic cases and with particular attention to how data quality in terms of resolution and noise affects the effectiveness of the approach.

Key words: Hydrogeophysics; Electrical resistivity tomography (ERT); Inverse theory; Neural networks, fuzzy logic; Numerical modelling.

1 INTRODUCTION

In geoscience, the electrical resistivity tomography (ERT) is considered to be one of the most reliable and popular geophysical methods for imaging the shallow subsurface, as shown by the large number of applications in the research and engineering fields on various issues such as mapping geological structures; detecting faults, karstic sinkholes and cavities (Cardarelli *et al.* 2006; Youssef *et al.* 2012; Billi *et al.* 2016); delimiting contaminated soil (LaBrecque *et al.* 1996; Werkema Jr *et al.* 2003); and the localization of buried archaeological remains (Negri *et al.* 2008; Cardarelli & Di Filippo 2009).

The ERT method consists in interpreting a set of apparent electrical resistivity data calculated from measurements of potential difference over an arrangement of potential and current injection electrodes. The interpretation process is based on the

implementation of an optimization code to find an electrical resistivity model, of which the simulated apparent resistivity matches the data set acquired over the field. This process is iterative and requires multiple numerical resolutions of the Poisson's equation before achieving convergence. This operation is called the inverse problem and formulated mathematically as a nonlinear, underdetermined and ill-posed optimization. The ERT inverse problem has been widely discussed in the literature with the use of various categories of deterministic, stochastic and global algorithms, and with different constraints and *prior* information. Deterministic methods such as the Gauss–Newton, quasi-Newton and conjugate gradients are the most commonly used schemes in the geoelectric surveys due to their rapid computation, in which the sensitivity matrix is computed at each iteration in the minimization of the objective function (Park & Van 1991; Loke & Barker 1996; Oldenburg & Li 1999; Kemna *et al.* 2000; Günther *et al.* 2006). However, the inverted

solution from these methods is a local minimum and strongly depends upon the initial model used to launch the algorithm. Stochastic methods represent an effective strategy to deal with the uncertainties of the inverse problem of ERT by sampling a set of models instead of a single solution. However, to reach the convergence state, the algorithms would have to test thousands of randomly proposed models, which slows down the computation (Jardani *et al.* 2013).

The global optimization algorithms have also been applied for predicting the best subsurface resistivity model by searching for global optima. Unlike deterministic methods, these algorithms do not depend on the initial model; among them are the particle swarm optimizations (Fernández-Martínez *et al.* 2010; Francisco *et al.* 2019), the simulated annealing (Sharma 2012; Sharma & Biswas 2013), genetic algorithms (Furman *et al.* 2004; Başokur & Akca 2011) and the deep learning with neural networks (NNs; El-Qady & Ushijima 2001; Feibo *et al.* 2018; Liu *et al.* 2020). The first three algorithms use an optimization concept to find a resistivity model suited to the measured data, while the NN reconstructs the soil model directly from the apparent resistivity data without applying an iterative optimization algorithm on these data. Indeed, the NN works as an end-to-end operator that approximates the inverse function by estimating the network parameters in the training process with a set of known data. Hornik *et al.* (1990) have shown that multilayer NNs can provide the accurate approximations of complex and highly nonlinear functions including the inverse operator by analysing statistically the relationship between the input and output parameters of training data. In these supervised deep learning algorithms, the preparation of training data represents a crucial step in the process, and in ERT application takes up most of the time because it involves the use of numerical tools to solve partial differential equations. However, the interpretation of geoelectric data with the predictive inverse operator takes only few seconds (Araya-Polo *et al.* 2017).

The earliest attempts at applying deep learning to ERT were made using conventional NNs in which each neuron is connected to other neurons forming fully connected layers. This architecture involves the learning of a large number of parameters during the training phase, which makes computation complex in the prediction of a large-scale and nonlinear function (El-Qady & Ushijima 2001). Thus, these early fully connected architectures were not well suited to handling tomography with high-dimensional input and limited computing resources.

However, the emergence of the new generation of powerful computer and deep learning (DL) architectures has increased considerably the popularity of the machine learning algorithms and brought about a radical change in the way we process and interpret geoscience data (Lary *et al.* 2016; Yang & Ma 2019). On the tomographic theory, several promising approaches inspired by recent techniques developed in the realms of image processing and medical imaging such as the convolutional neural networks (CNNs) have been published; for example, Puzyrev (2019) used the CNNs with a fully connected layer to map in a 3-D spatial distribution of electrical conductivity by training the low–medium-frequency electromagnetic data. Wei & Chen (2019) proposed the CNNs under U-Net architecture to predict the inverse operator linking the full electromagnetic wave scattering data and dielectric field. Li *et al.* (2020) applied the deep learning algorithm to deal with the processing of seismic data in order to image spatial heterogeneity of the seismic velocity of the ground. Also, in the seismic exploration, Araya-Polo *et al.* (2018) discussed the relevance of the deep learning algorithm to retrieve the velocity model of the subsurface

from seismic data. Liu *et al.* (2020) interpreted the apparent resistivity data acquired in the ERT survey by using the CNN method reshaped according to the U-Net architecture. This approach has been successfully validated on the theoretical cases in 2-D.

Indeed, CNN is a technique that arouses the interest of the geophysicist community, thanks to its capacity to process tomographic images in a reasonable time. The main idea behind this concept is the use of a convolution operator in which several filters are applied to extract the local characteristics of each data set. Each filter identifies a feature type by scanning it over the input image that hence results in the feature map with low resolution, which drops drastically the number of the parameters to train (LeCun *et al.* 1998). The process is sequentially performed through multiple convolutional layers; the choice of their number is defined according to the complexity and the nature of tasks. For the tomographic inverse problem, the CNN with encoder–decoder architecture has gained popularity and has been tested on various cases such as the positron of emission tomography (Häggström *et al.* 2019), the full electromagnetic wave scattering (Wei & Chen 2019), the ERT (Liu *et al.* 2020) and the seismic processing (Apolinario *et al.* 2019; Moseley *et al.* 2020).

These architectures were initially introduced for image segmentation in identifying the typology of the classified objects on images such as U-Net and SegNet architectures (Badrinarayanan *et al.* 2017; Yao *et al.* 2018). These structures process the input image through two stages: the first is the encoder operation in which the learning is effectuated with the downsampling feature maps obtained with the successive convolutional, nonlinear activation and pooling methods. The second is the decoder network regarded as a mirror of encoder operator that is dedicated to recover the initial resolution condensed in the encoder via the use of max-pooling indices derived from the pooling layers at the encoder to upsample feature maps (Badrinarayanan *et al.* 2017). Such architecture permits to have a reliable generalization in conserving resolution and details of the image and reducing the number of the trainable parameters (Badrinarayanan *et al.* 2017). In this paper, we explore for the first time, according to our knowledge, the use of a CNN method with SegNet architecture to reconstruct in 3-D the spatial variability of the electrical resistivity of the subsurface.

The paper is structured as follows: the theoretical background is first presented, which includes the generation of training data, resistivity acquisition and the proposed network architecture. The network is then applied to predict a geostatistical ground model in Section 3. Several representative examples are illustrated to evaluate the effectiveness of the network. Impacts of potential errors due to the resolution and noise in the measurement data are also discussed. The trained network is later extended for monitoring the evolution of a slat plume in subsurface. Finally, we summarize our work in the conclusion section.

2 THEORETICAL BACKGROUND

The methodology of this approach consists of three essential parts: First, the preparation of the feeding data to form the network using theoretical models derived from the numerical resolution of the Poisson's equation. Second, the representation of the measured apparent resistivity data as a 3-D apparent resistivity field to form the input image of the deep learning network. Third, the determination of the parameters of the 3-D encoder–decoder network in the learning phase in order to establish a link between the apparent electrical resistivity field and the real resistivity distribution assigned to the geological model.

2.1 Generation of training data

In general, the supervised machine-learning algorithm requires a large amount of data to feed the learning networks for providing an accurate generalization, which makes data assembly a crucial step in the process. In ERT, the acquisition of the input and output data, which in this case are the apparent resistivity data recorded at the ground surface and the subsurface distribution of electrical resistivity, could only be obtained theoretically due to the difficulty of deriving them in the field. For this reason, we use the forward operator ERT that allows to link the 3-D underground distribution of electrical resistivity with the apparent electrical resistivity measurements collected by the potential electrodes via the resolution of the Poisson's equation:

$$\nabla \cdot \left(-\frac{1}{\rho} \nabla \phi \right) = I_0 [\delta(\mathbf{r} - \mathbf{r}^+) - \delta(\mathbf{r} - \mathbf{r}^-)], \quad (1)$$

where ρ (Ω m) denotes the electrical resistivity, ϕ (V) is the electric potential at location \mathbf{r} in response to the current injected I_0 (A) between two electrodes placed at \mathbf{r}^- and \mathbf{r}^+ , and δ is the Dirac distribution.

For a given spatial distribution of electrical resistivity and boundary conditions, the electric potential field is calculated numerically at any point in the domain, including the surface, by using finite element or finite difference methods. Then the voltage data determined at the electrode positions are converted to an apparent resistivity, ρ_a (Ω m), using the following formula (Koefoed 1979):

$$\rho_a = k \frac{\Delta V}{I_0}, \quad (2)$$

where k (m) is the geometric coefficient that depends on the spatial configuration of the current and potential dipoles, ΔV (V) is the voltage difference between the potential electrodes.

In this paper, we solve the forward problem for each 3-D distribution of electrical resistivity to feed the training networks. The assembly of these training data is obtained thanks to a geostatistical generator of subsurface resistivity and the integration of these models in the forward problem to determine the apparent resistivity field for a given acquisition configuration. In the end, the training data consist of the output subsurface resistivity models and their corresponding input apparent resistivity fields. We reserve a subset of these unused data in the training algorithm to validate the effectiveness of the approach.

2.2 Construction of apparent resistivity sections

In this study, we aim to cover in 3-D the spatial heterogeneity of resistivity with several 2-D profiles; this strategy effectively reduces the cost and complexity of data acquisition and offers the possibility of processing the data in 2-D (Loke *et al.* 2013). The apparent resistivity data recorded with the electrodes are then used to form a pseudo-section in 2-D by positioning each measurement at a given location, where the pseudo-depth can be theoretically derived by the formulation of Roy & Apparao (1971) or simply using the distance between the electrodes. At the end, we interpolate all these 2-D apparent resistivity sections to form a 3-D pseudo-map of apparent resistivity using the nearest-neighbour interpolation method. The nearest interpolation requires a lower computational cost and avoids extreme values that may be related to a abrupt change in the medium. This interpolation brings the input images to the same size as the output images before starting the training processes of the encoder-decoder networks (see Fig. 1). It is possible to adopt the same

strategy in the construction of 3-D pseudo-apparent resistivity maps with other common or uncommon resistivity acquisition devices with electrodes placed on the surface and in boreholes with parallel or radial schemas.

2.3 CNN architecture

We recall that the resolution of the inverse process aims to find a model of subsurface resistivity \mathbf{m} that provides a good numerical match with the apparent resistivity data acquired at the ground surface \mathbf{d} . However, this model is not unique due to the scarcity and uncertainties of the data, making it impossible to cover all the heterogeneities of the studied environment. Most commonly, this issue is handled with the use of a regularization term. In machine learning, the inverse problem is however reformulated as a black box model in which output and input are linked by a large number of parameters defined on multiple NNs. These parameters are tuned in an optimization algorithm to suit the training data. In fact, machine learning is designed to approach a universal inverse operator for a group of models that have features in common. Then, this inverse operator can be expressed in function of training data $\{\mathbf{d}, \mathbf{m}\}$ and deep learning parameters Θ :

$$\mathbf{m} = f_{\text{inv}}(\mathbf{d}, \Theta). \quad (3)$$

This inverse operator can be approached by inferring deep learning parameters Θ through the minimization of this objective function:

$$\Theta = \arg \min \left\{ \sum_{i=1}^N \|m_i - f_{\text{inv}}(d_i, \Theta)\| \right\}, \quad (4)$$

where $\{\mathbf{d}, \mathbf{m}\}$ refers to the pair of apparent resistivity data and their corresponding subsurface resistivity. N is the number of training data used to predict the inverse operator f_{inv} that is parametrized via Θ .

The main objective of this work is to implement an NN architecture that can efficiently provide an inverse operator with limited computing resources. This efficiency is often relevant to complexity and refinement of the network adopted. For this reason, we aim to design a complex architecture based on the SegNet since this structured network gained recognition by its computational performance and efficiency among number of architectures. The SegNet architecture that has been successfully applied to process semantic segmentation of images in accurately retrieving and recognizing the spatial forms of learned objects (car, road, etc.; Badrinarayanan *et al.* 2017). Thus, the original architecture was conceptualized for a classification purpose in which the training process is conducted sequentially with the encoder and decoder operators. However, in our case, it is generally challenging to make an effective prediction about the labelling from *prior* information in the real field.

This is why we aim at a regression network in which the resulting image is represented by quantitative levels instead of a classification result. This such quantitative output conserves more detailed information and probably assists more in post-processing analysis. The architecture consists of an encoder network and a corresponding decoder network with a regression layer at the end as illustrated in Fig. 2. Each encoder and decoder network composes 21 layers.

The encoder network has three encoders; each encoder contains a convolutional layer, where a set of 64 filters of 3×3 convolutes the input image to bring out local feature maps. On these feature maps, a batch normalization and ReLU layer operate successively to introduce nonlinearity into the process. The result of the previous

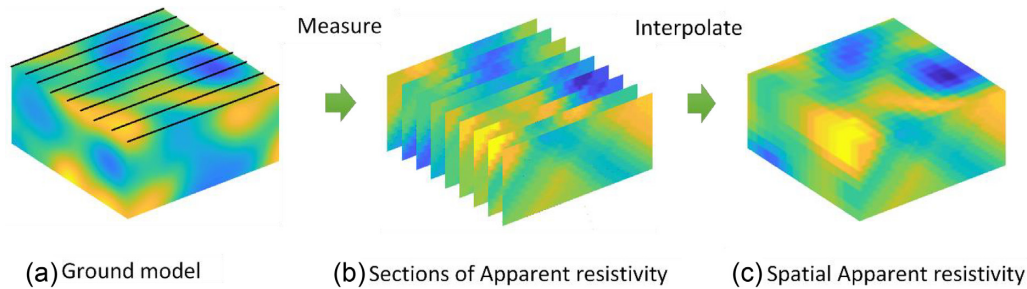


Figure 1. Description in image of the method applied in the genesis of the training data. Panel (a) represents the model of the true subsurface resistivity distribution used in the forward operator to derive a set of apparent resistivity sections acquired on each profile (b), which are interpolated in the next step to construct a 3-D image of the apparent resistivity (c).

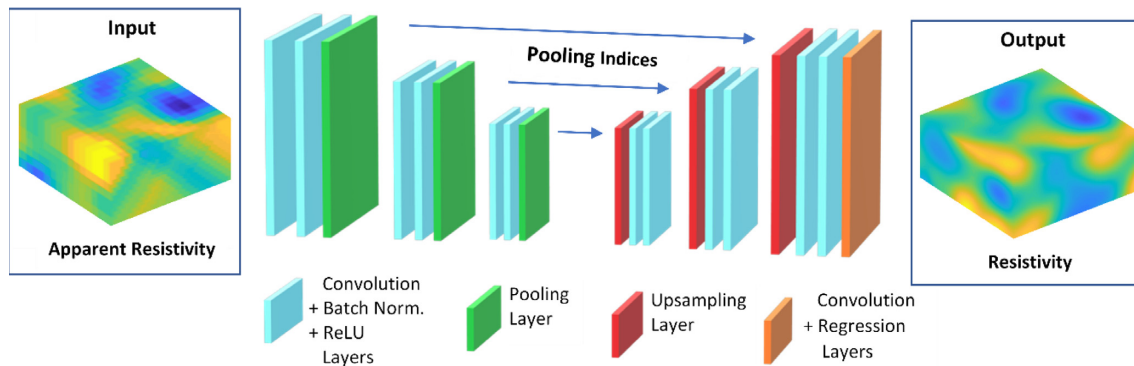


Figure 2. Illustration of the SegNet architecture proposed. The architecture is developed with three-level encoder–decoder structure, including 45 layers in total; the classification layer at the output is replaced by a regression layer to recover the electrical resistivity map. The apparent resistivity from measurements feeds the input data for the network (modified from Badrinarayanan *et al.* 2017).

layer in turn is processed by a max-pooling layer of 2×2 window and stride 2, which induces a loss of spatial resolution. Therefore, the feature maps resulting from each encoder operator are switched to its respective decoder (the pooling indices) to counterbalance this loss of resolution. In the decoder network, the process starts with an upsampling layer in which the feature maps retrieved from the encoder are concatenated to improve the resolution. Then the results from the previous layer will be successively processed by the convolution, batch normalization and ReLU operations (Badrinarayanan *et al.* 2017). The final pixel-wise classification in the original architecture designed for object classification is replaced by a convolutional layer followed by a regression layer to operate the regression task in our problem.

As mentioned above, in the decoding process, SegNet uses maximum pooling indices to communicate information from the encoder to its corresponding decoder, thus saving memory. Whereas in UNet, entire feature maps of the encoder are transferred to the decoders, which actually needs more memory. For a comparable configuration in this test, the UNet architecture requires about eight million parameters to be learned, but SegNet only works with half a million, which is a fraction of the cost. For more details on SegNet's ability to perform complex tasks compared to other networks, we invite readers to see Badrinarayanan *et al.* (2017).

3 APPLICATIONS

3.1 Geostatistical ground model

In this paper, we focus on the reconstruction of the subsurface resistivity models having in common certain statistical properties.

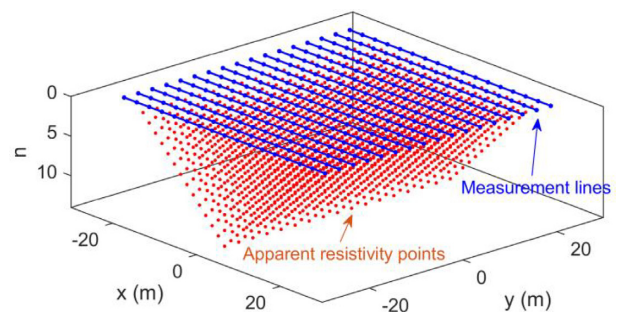


Figure 3. Spatial distribution of the 17 profiles used in this study where the electrodes are regularly separated by 3 m (blue dots). For each profile, we reconstruct a section of the apparent resistivity with 105 measurements (red dots) acquired using the dipole–dipole array where the n is data level.

The logarithms of these models are generated geostatistically with a Gaussian distribution with variable properties. In this respect, the logarithm resistivity is drawn in a random manner from the intervals where the mean is from one to three orders, the range is from 5 to 100 m and the rotate angle of anisotropic is from 0° to 360° . We note that the degree of heterogeneity of the resistivity fields generated is variable in all three directions, and may reflect the complexity encountered in some real field conditions. The models vary from 1 to 5 orders with resistivities ranging from 3 to 10^5 .

The 3-D spatial heterogeneity of the resistivity is determined from multiple 2-D profiles aligned in parallel along the y -direction (see Fig. 3). Each profile is composed of 17 electrodes separated by 3 m on which 105 measurements of apparent resistivity were collected with a quadruple dipole–dipole. The same operation was

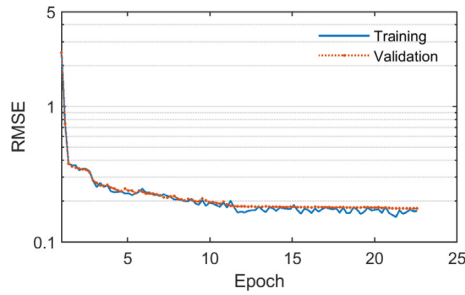


Figure 4. Results of the training process with 30 000 samples for training and 5000 samples for validation, the training is stopped manually to avoid overfitting.

repeated on the rest of the profiles, which are 3 m apart, for 1.785 measurements over the entire set of profiles.

3.2 Learning and validation of the NN

A total of 45 000 samples are collected, of which 35 000 are assigned to network learning including 30 000 models for training and other 5000 models are selected randomly for validation during training process. The rest (10 000 samples) are reserved for testing after the training operation (named as Case C0). Each subsurface resistivity model contains $32 \times 32 \times 32$ voxel grids, and the same size is used in the reconstruction by interpolation of the apparent resistivity field. These models take up 10 gigabytes of memory. As discussed in the previous section, the deep learning training is done through a modified version of the SegNet with an encoder–decoder network and fully connected layer placed at the end to perform a regression task between input and output images. The network is performed with the ADAM optimizer implemented in Matlab software on a workstation with NVIDIA QUADRO K2200 and 64G RAM. The training task is performed with a batch size of 64, constant learning rate of 0.01 and occupies up to 30 gigabyte memory. Typically, the learning convergence lasts 1.5 hr in 23 epochs, the training progress results are plotted in Fig. 4.

Once the training is complete, we examine the generalization efficiency using 10 000 unseen models, and this test takes only a few minutes to yield the response in terms of subsurface resistivity models.

For the metric evaluations of the quality of our predictions, we employ the coefficient of determination (R^2) and the root-mean-square error (RMSE) defined as follows:

$$R^2 = 1 - \frac{SSR}{SST}, \quad (5)$$

$$RMSE = \sqrt{\frac{\sum_1^n (s_{\text{pred}} - s_{\text{true}})^2}{n}}, \quad (6)$$

where $SSR = \sum_1^n (s_{\text{pred}} - s_{\text{true}})^2$; $SST = \sum_1^n (s_{\text{true}} - \bar{s}_{\text{true}})^2$, with s_{pred} and s_{true} denote the predicted and true logarithm of resistivities $s = \log_{10} \rho$, respectively; \bar{s}_{true} is the mean of s_{true} ; and n denotes the number of voxel grids.

In general, most of these models are well predicted with an average correlation with the real resistivity and the predicted fields is about 0.16 for RMSE and 0.85 for R^2 ; RMSE of testing sets is closed to RMSEs of training and validation data sets (0.15). To better evaluate the results obtained, we have chosen to show three

representative reconstructions of all validation models with different degrees of heterogeneity. Fig. 5(a) corresponds to a model with flattened variability of electrical resistivity that has been accurately reconstructed. While the model shown in Fig. 5(d) presents an intermediate heterogeneity, particularly in horizontal directions and its prediction reproduces the main original features. However, in the last example, Fig. 5(g), the model is characterized by the presence of spatially narrow anomalies which make the model a bit more complex; its reconstruction therefore is less perfect than the previous examples but still faithful in capturing the main anomalies.

In general, the discussed examples show that the network can adequately capture the magnitude and variability of the target subsurface resistivity model. Some difficulties arise in particular on the boundaries of the model due to the loss of resolution, which is also the case with conventional inversion methods. The accuracy of network estimation can also be conditioned by the degree of complexity of the predicted model. A model with less variability is an easier model to predict. As noted above in Section 1, deep learning approach relies on data, and therefore the nature and quantity of the input data set used in the training phase will control the accuracy of prediction. Since these data result from the interpolation of the 17 pseudo-sections of apparent resistivity, some aspects in the reconstruction will not be detectable by the electrode arrangements and others will be hidden in the interpolation process. Further discussion on the influence of the spatial resolution of the input image will be addressed in Section 3.2.3.

3.2.1 Influence of data set size in the training process

The deep learning approach is founded on a statistical approach in which the quantity and quality of the training data set will be decisive in the performance of the trained network. In this subsection, we first address the effect of the volume of training data by examining the quality of forecasts with different numbers of data used in the training task: 15 000, 20 000, 30 000 samples. The results are summarized in Table 1, which shows that the accuracy of predictions increases with the size of the training data. In fact, the reliability of prediction models is controlled by the amount of input data, so such dependency can lead to a high computational cost for database generation.

The choice of the data volume required in the training phase depends both on the heterogeneity of the models to be predicted and on the degree of precision desired. As shown in Figs 6(a)–(c), a resistivity model with moderate variability was reconstructed with reasonable accuracy, even with a limited number of training data. However, for a more complex heterogeneity model as shown in Figs 6(d)–(f) and (g)–(i), increasing the size of data allows a better reconstruction of the heterogeneities. As this accuracy also depends on the acquisition protocol and a large number of measurements allows a better coverage of heterogeneities even with a limited number of training data.

3.2.2 Influence of data resolution

We devote this section for analysing the impact of training image resolution on the quality of predictions. To do so, we preserve the size of the input image $32 \times 32 \times 32$; however, we reduce the quality of the information carried by the input data by using only nine pseudo-sections in the construction by interpolation of the input images (see Fig. 7). The prediction of the inverse operator derived with few profiles is still efficient to map the main heterogeneities

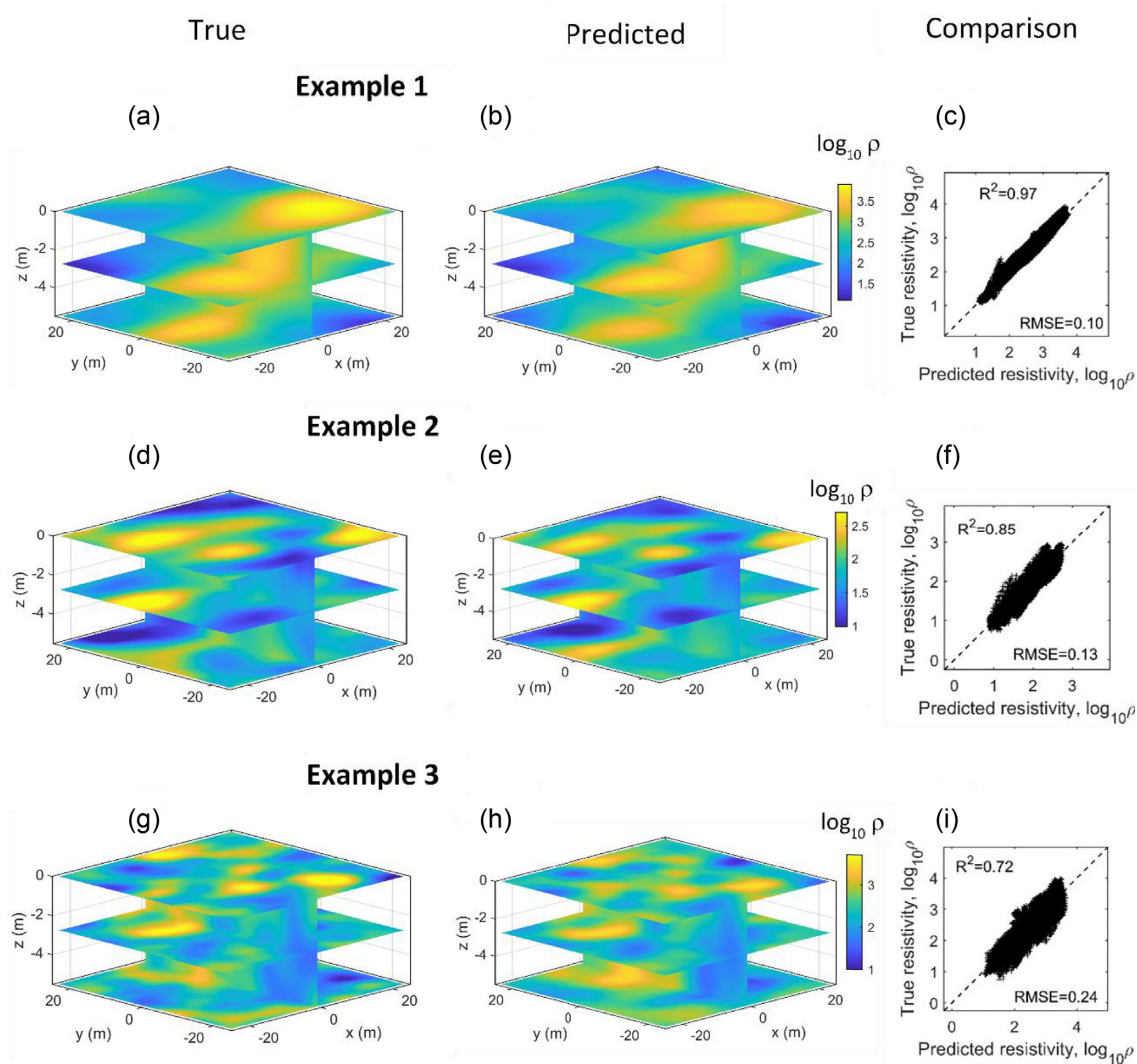


Figure 5. Examples of inverted tomography employing the CNN network (Case C0). By row, from top to bottom: three examples show three representative reconstructions of all validation models with different degrees of heterogeneity. By column, from left to right: true resistivity–predicted resistivity–comparison between them. The network reproduces properly both value and heterogeneity in the medium.

Table 1. Correlation coefficient (R^2) between the true and inverted resistivity fields for different size of training data set.

	Size of training data		
	15000	20000	30000
Example 1	0.92	0.96	0.97
Example 2	0.71	0.77	0.85
Example 3	0.52	0.69	0.72
Overall (10 000 samples)	0.70	0.81	0.85

of the validation models with a correlation coefficient average of $R^2 = 0.79$. The models illustrated for the discussion in the previous section are also well-characterized, and this is due to the nature of the heterogeneities of the models, which can still be covered with nine profiles and therefore unaffected by the reduction in data. However, if we analyse only the models with a high level of heterogeneity, the quality of the reconstruction is impaired by the lack of data.

In the second test, we further reduce the number of profiles used in the survey to five profiles and the results are summarized in Table 2. As in the previous cases in terms of data construction and training, the predictions of the validation models revealed a degradation of the characterization. Thus, as for the rest of the inversion techniques, the quality of tomography depends on how well the measurements can cover the heterogeneity of the subsurface.

3.2.3 Influence of noise in data feeding

To expand the discussion of the factors influencing inversion performance through further learning, we explore below the effects of noise contamination on feed data. Here, we contaminate the input data by adding 25 and 50 per cent Gaussian noise to the measured apparent resistivity. Analysis of the prediction results on the validation sets shows that the mean R^2 determination coefficients are 0.77 and 0.49, respectively, which represents a reasonable degradation of accuracy (see Fig. 8 for representative examples). However, the

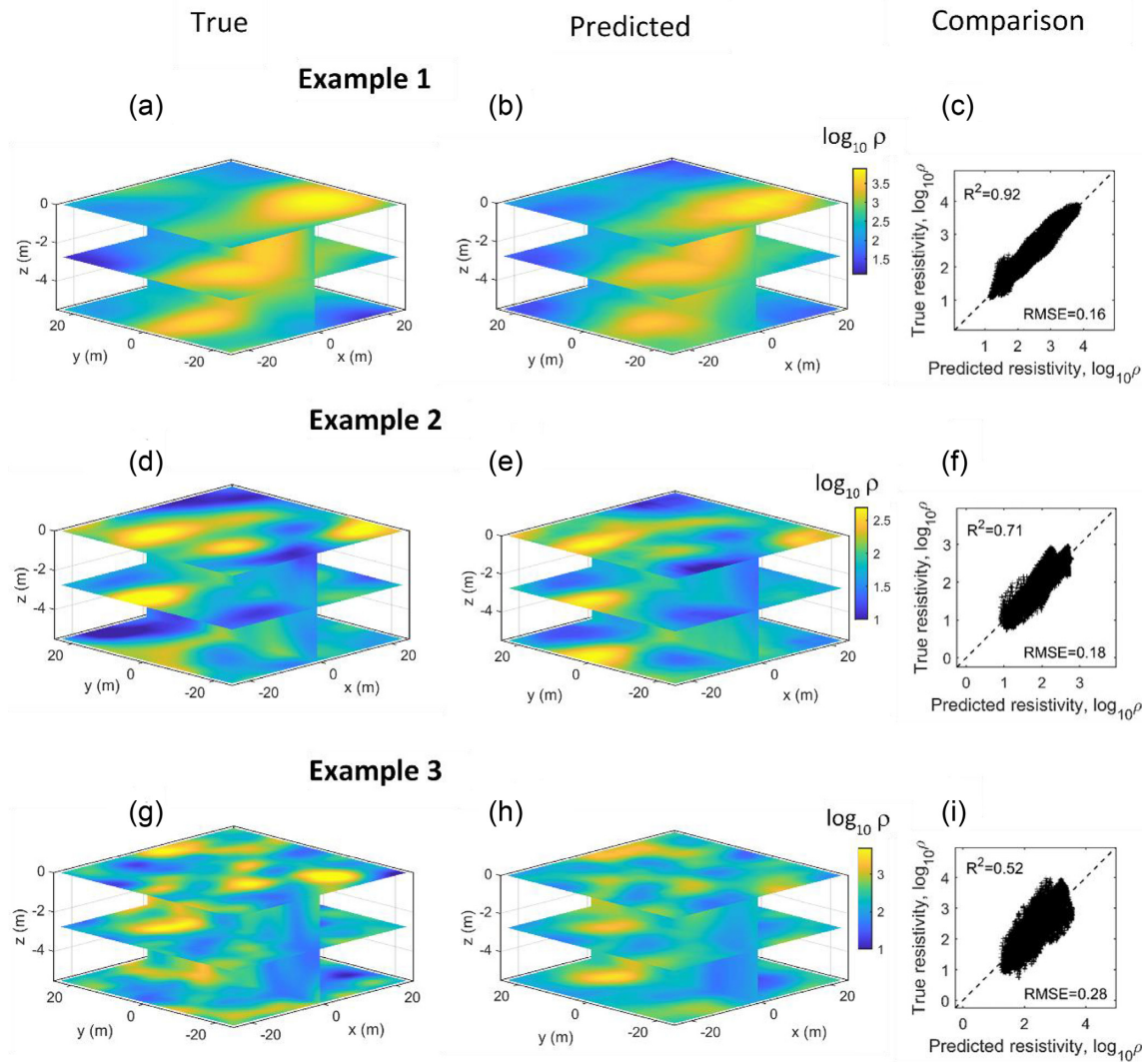


Figure 6. Predictions from the network using lower training data set of 15 000 samples (named as Case C1). The accuracy is generally lower than in Fig. 5 obtained with 30 000 samples.

overall noise impact is still minimal if it is not modifying the characteristics of the models. Similar conclusions are also noted in other studies using comparable network architectures, such as Inversion-Net (Wu & Lin 2018) or U-Net (Wei & Chen 2019). This can be attributed to the concept of CNN networks, where data processing is based on the analysis of spatial models rather than individual data points.

As an outline to clarify the result of discussions in this test, we summarize all comparisons between the true and predicted resistivity tomographies in Table 2; the details of histograms for overall testing of 10 000 samples are compared in Fig. 9. In the next section, we will adopt this trained network in a synthetic monitoring test to verify the applicability and performance of the proposed architecture.

3.2.4 Influence of the features of training models

In order to have a better vision of the efficiency and limits of the approach, this time we apply the network formed in the previous test

(Case C0) with geostatistical models to interpret apparent resistivity sections acquired on models with different heterogeneities than the training models. These models are characterized by sharp horizontal or vertical discontinuities. The first model consists of a ground model with two horizontal layers: an upper layer of 1000 Ωm and a lower layer of 10 Ωm . However, second and third models have three and four vertical layers, respectively (Fig. 10). These models are used to create pseudo-sections of apparent resistivity with the same acquisition protocol developed in the first test (17 profiles). These sections were then introduced as input data into the network to analyse its quality of predictions on models with characteristics that were not seen in the training data sets. The predicted models are presented in Fig. 10 where the horizontality and verticality of the structures have been identified, but are nevertheless perturbed by the presence of certain artefacts that do not exist in the real models (e.g. the appearance of a conductive zone on the surface of the first model). The comparison between the resistivity values of the real model and the predicted model shows that the reconstruction is less satisfactory. This was also confirmed by the coefficient of determi-

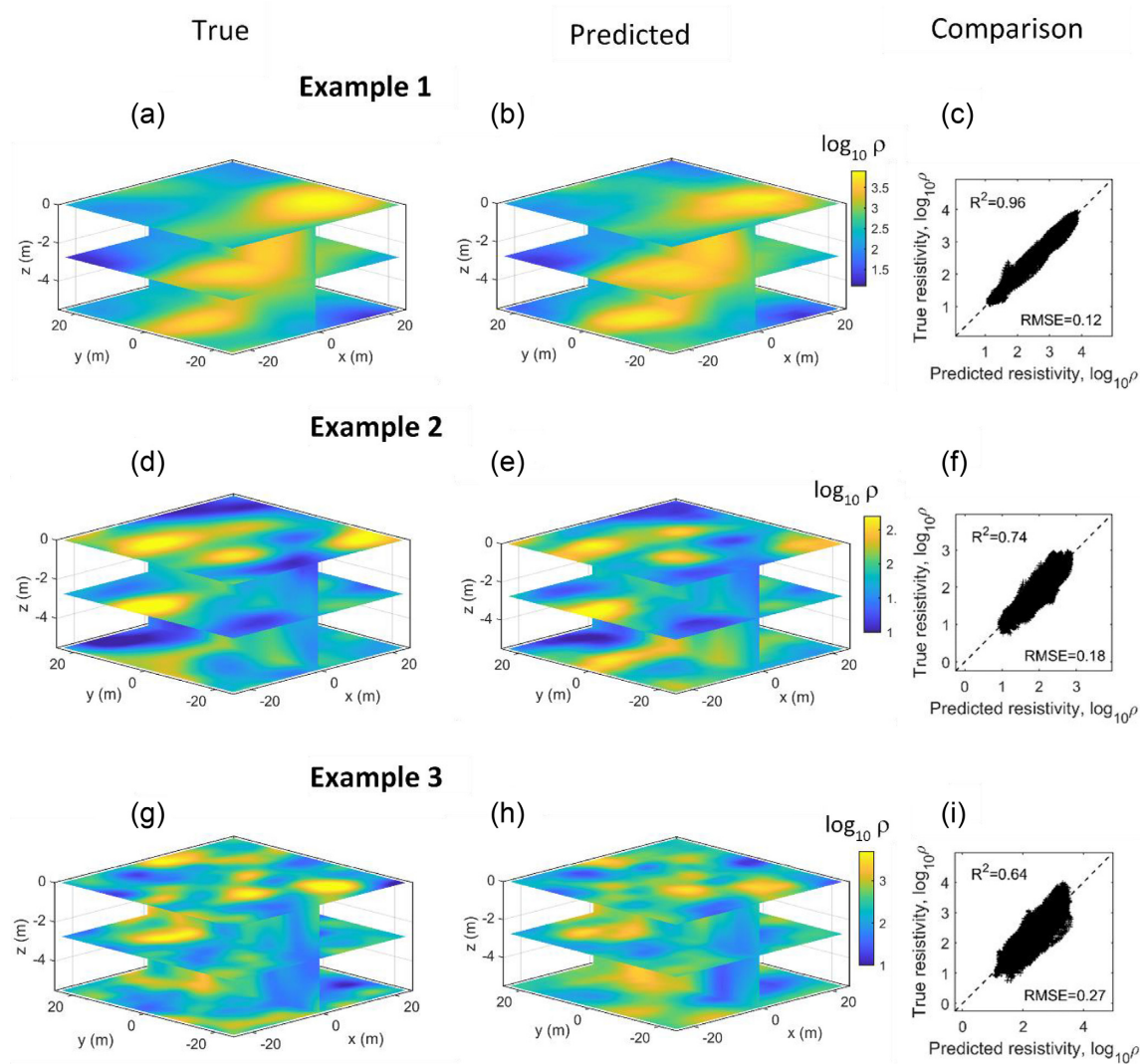


Figure 7. Predictions with lower-resolution data interpolated from nine sections (instead of whole 17 sections)—named as Case C2. The results show a different decrement in accuracy comparing to Fig. 5 which depends on the heterogeneity of samples. The quality of the reconstruction is impaired by the lack of data.

Table 2. Correlation coefficient (R^2) between the true and inverted resistivity field for different interferences. By row, results for typical examples are shown for details and overall testing data of 10 000 samples. By column, results are for uncontaminated testing data, and contaminated testing data of lower resolution or with noise.

	Uncontaminated data (17 sections)	Lower resolution		With noise	
		9 sections	5 sections	25 per cent	50 per cent
Example 1	0.97	0.96	0.73	0.95	0.85
Example 2	0.85	0.74	0.44	0.78	0.38
Example 3	0.72	0.64	0.17	0.64	0.29
Overall (10 000 samples)	0.85	0.79	0.60	0.77	0.49

nation and the RMSE calculated between the apparent resistivity data derived from the real and the predicted models ($R^2 < 0.35$ and $RMSE > 1$), which reveal that the optimization has failed. This is not surprising, since the inversion with the deep learning algorithm cannot be applied to models with characteristics different from those used in learning. Therefore, the results of inversion with deep learning codes also depend on the prior information used in the generation of training models, as is the case with other traditional inversion methods.

3.3 Monitoring of a salt plume migration

In this section, we apply the deep learning approach to interpreting ERT data acquired in tracking the movement of a conductive body, such as saline plume. The process of interpreting this type of data with conventional inverse methods is a computationally expensive task because many ERT data have to be inverted for each time. Whereas deep learning requires only one training operation to interpret the entire data collected at different moments in time. This

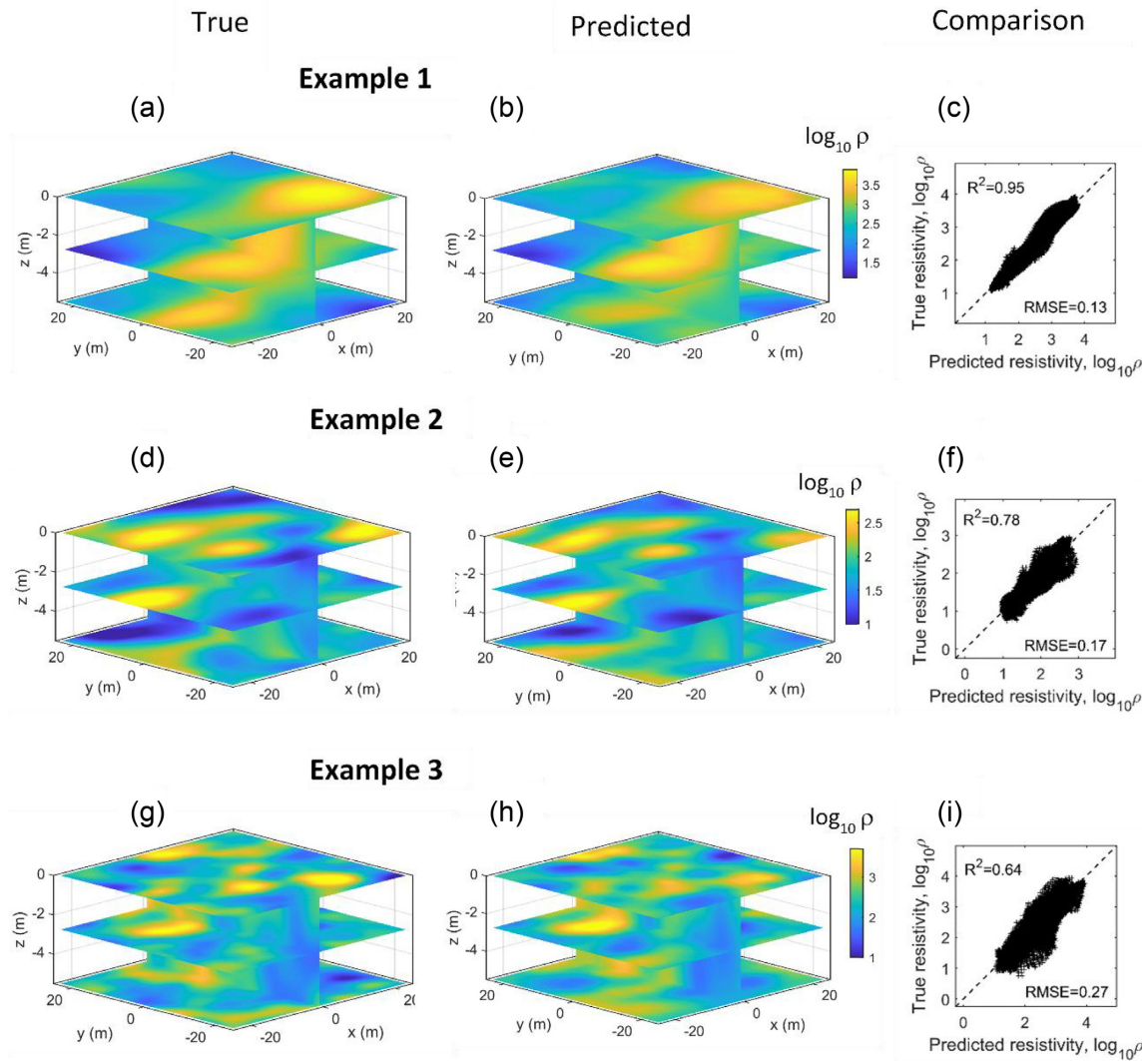


Figure 8. Examples of predictions from the apparent resistivity contaminated with noise 25 per cent with the same examples as shown in Fig. 5 (named as Case C3). The overall noise impact is reasonable as it is not modifying the characteristics of the models when data processing is based on the analysis of spatial models rather than individual data points.

numerical case is designed to monitor the movement of a salt tracer by remotely measuring at the surface the change in resistivity in the subsurface. A salt tracer was injected via a well (with a constant rate of $5 \text{ m}^3 \text{ d}^{-1} \text{ m}^{-1}$ and a concentration of 100 Mol m^{-3}) placed on a homogeneous aquifer with a hydraulic conductivity of 10^{-4} m s^{-1} and a thickness of 30 m, the background flow associated with a hydraulic gradient five per cent along the y axis (see Fig. 11).

The spatial dispersion of the salt plume is determined numerically by solving in coupling ways and in transient mode the groundwater flow and transport equations. The consequence of this conductive plume on the modifications of the electrical resistivity of the aquifer is derived according to model of Waxman and Smits (1968) as

$$\sigma = \frac{\sigma_f}{F} + \sigma_s, \quad (6)$$

$$\sigma_f = a c (\beta_+ + \beta_-) e, \quad (7)$$

where $F = \theta^{-m}$ ($\theta = 0.3$ is the porosity and exponent parameter $m = 1.5$). σ_s is the surface conductivity assumed homogenous with a value of 0.005 S m^{-1} , σ_f (S m^{-1}) is

the poral electrical conductivity that linearly depends on the solute concentration c . e is the elementary charge ($e = 1.6 \times 10^{-19} \text{ C}$); $\beta_+ = 5 \times 10^{-8} \text{ m}^2 \text{ s}^{-1} \text{ V}^{-1}$ and $\beta_- = 7 \times 10^{-8} \text{ m}^2 \text{ s}^{-1} \text{ V}^{-1}$ are the mobility of the cations and anions in the pore water at 25°C , respectively, the coefficient a converts the unit of concentration from kg m^{-3} to Mol m^{-3} .

To monitor the salt plume, we keep the experimental scheme used in the previous cases where the electrodes are separated by 3 m and placed on the 17 uniformly spaced profiles. Apparent resistivity data are acquired at different periods in time [1, 5 and 50 d] to delineate the spatio-temporal dynamics of the plume that follows a Gaussian distribution centring on the injected well due to the homogeneity of the aquifer. We re-apply the network built in the previous section to identify the heterogeneity due to the movement of the salt plume. The reconstructions obtained at different periods are shown in Fig. 12, where the correlations between the real and predicted fields are in good agreement. For more details, the quality of the prediction has a certain dependency on the geometry of the anomaly. In fact, it is difficult to delineate a small anomaly from surface electrodes as shown in Figs 11(a)–(c) for the results of the

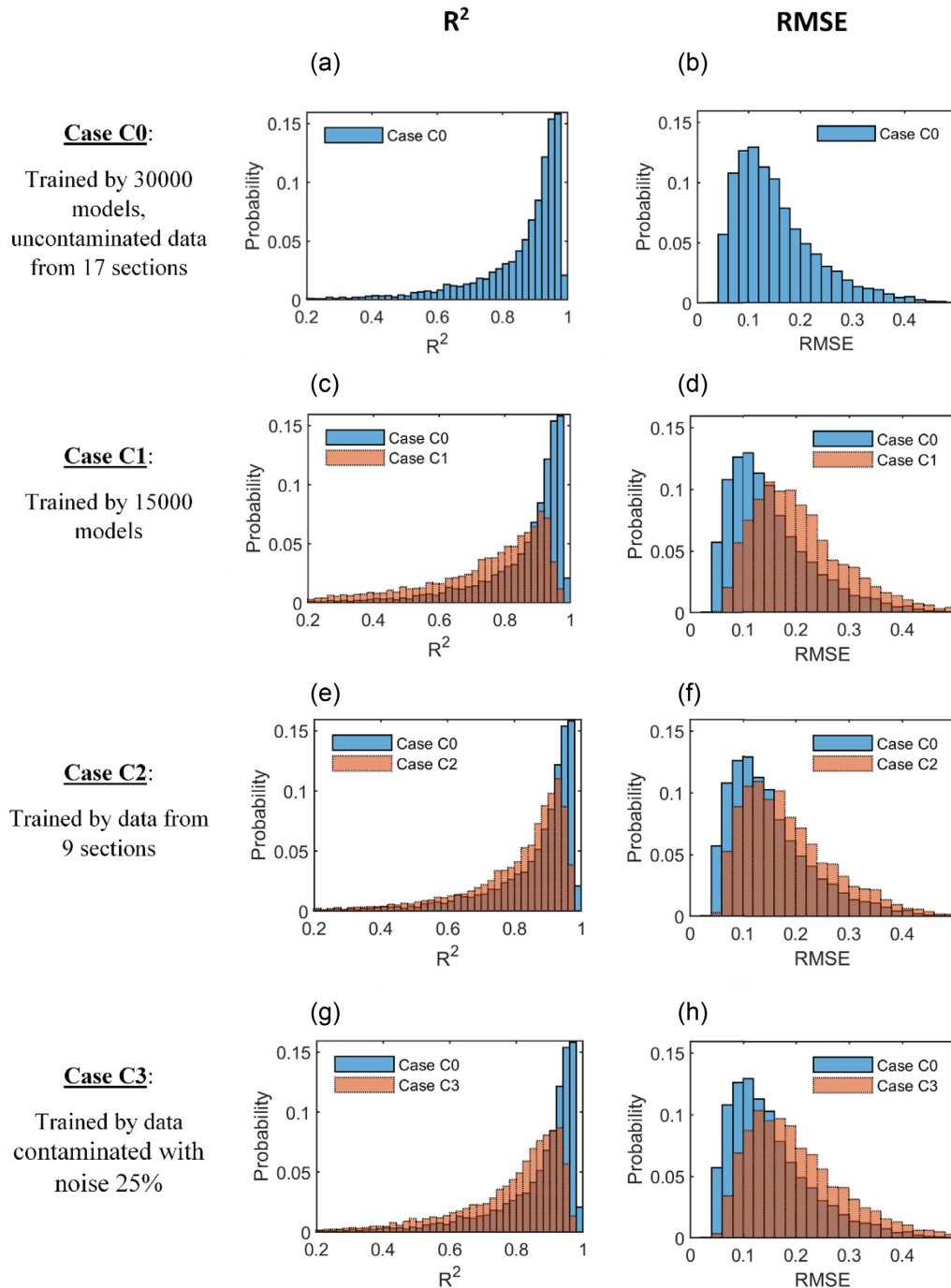


Figure 9. Histogram of testing data set containing 10 000 samples for different interferences from top to bottom. From left to right, details are for names of interferences, R^2 and RMSE between the true and inverted logarithm of resistivity.

one-day injection. On the whole, the architecture used can accurately infer the shape of the salt plume at different times without too much effort, but this study is limited to homogeneous aquifers. For very heterogeneous aquifers with complex preferential pathways, the training models have to be generated again to include these characteristics.

4 CONCLUSION

In geophysics, the processing of data acquired in a non-intrusive manner at the surface by inversion tools remains a complex and

difficult task due to the ambiguity of interpretation linked to the non-uniqueness of the solutions obtained with such approaches. Most often, these inversion tools are based on deterministic algorithms that aim at finding local minima, which are highly dependent on the initial models. In this paper, we propose a new approach to the 3-D inversion of electrical resistivity that is based on the concept of deep learning. This approach has been successfully validated on heterogeneous models generated by a geostatistical method. It has also been applied to monitor the migration of a salt plume over time by inverting the apparent resistivity acquired during injection of salt into a shallow aquifer.

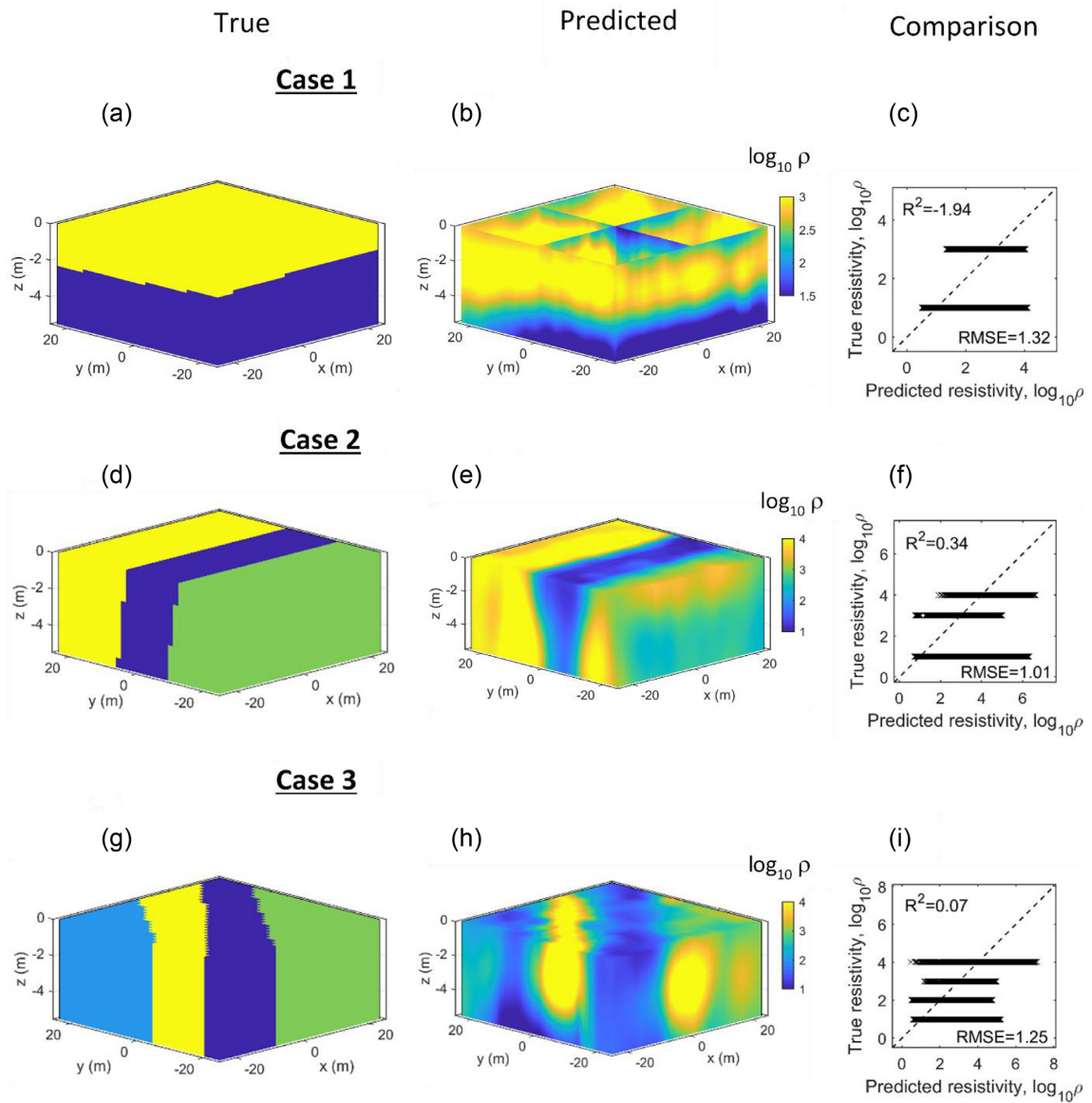


Figure 10. The application of a network formed with models generated with geostatistical constraint on models with vertical and horizontal layers. By row, from top to bottom: three examples show two-layer, three-layer and four-layer structures with various ranges of resistivity. By column, from left to right: true resistivity, predicted resistivity and comparison between them. The predictions obtained with the network do not clearly reproduce the geometry of the layers. The accuracy of the reconstruction with DL algorithm therefore depends on the nature of the heterogeneities chosen for the learning models.

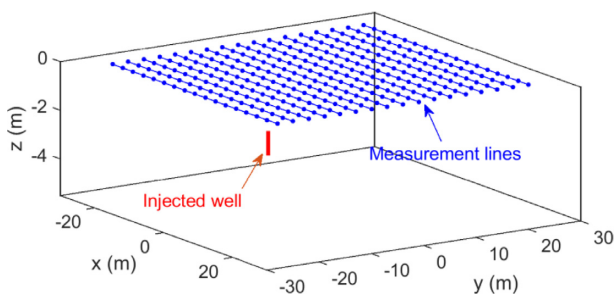


Figure 11. Setting of the injection test with the monitoring electrodes (blue points) and the injected well (red line) which locates in the middle beneath of the measurement electrodes.

In this deep learning tool, the training data are processed via CNNs that are organized on three-level encoder/decoder networks with a regression layer placed at the end of the decoder network to perform the regression calculation. The encoder network uses multiple convolutional, nonlinear activation and max-pooling operations to highlight key local features on the input images (apparent resistivity maps). The decoder network is designated in way to compensate the loss of the resolution occurred in the encoder network by employing the max-polling indices recovered from the precedent network for upsampling operation. The architecture is simple to implement and sufficient to approximate a highly nonlinear inverse operator in few hours of training. For the training step, a massive data set has been numerically generated by using the Poisson's equation to feed the NNs. In this paper, the feeding data contain 45 000 models that overcome the overfitting problem but they require tens of gigabyte memory. However, the training phase is executable on a

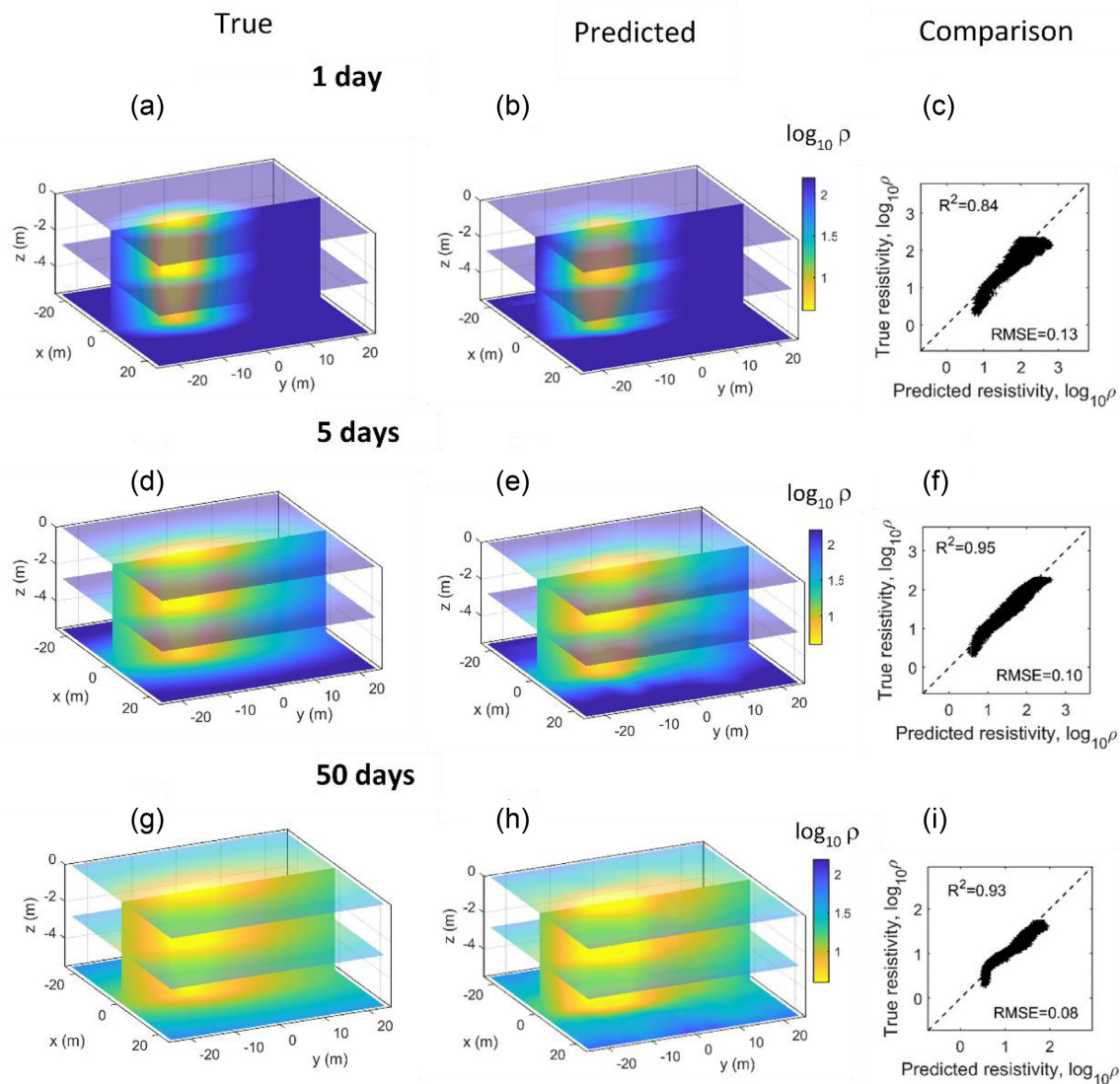


Figure 12. Evolution of the electrical resistivity in the subsurface due to a saline injection. By column, from left to right: true resistivity, predicted resistivity and comparison between them. By row, from top to bottom: data at 1, 5 and 50 d. The inverted results can reproduce both shape and value of the evolving field.

single GPU in a few hours, and prediction via the inverse operator will only take a few milliseconds. Thus, the effort remains on data generation.

The various analyses carried out on validation samples allow us to conclude that the characteristics of the training models have a considerable influence on the quality and nature of the predicted models. A network formed with models generated according to geostatistical characteristics can only be applied to provide models of the same nature as the training models. Therefore, it is essential that the characteristics of these training models in terms of spatial distribution and degree of variability be selected on the basis of prior information derived from geological studies (e.g. analysis of drilling data) or other geophysical, seismic or electromagnetic studies. For example, if the ERT survey is conducted over an area where the geological layers are tabular, the training models must be designed to contain this type of information, and the use of a formed network with a geostatistical distribution as described in our case may provide impertinent results. In case of doubt about the choice of the prior model, for instance, between two scenarios of geological structures, in this case the learning must be done with both models

simultaneously. However, such a strategy requires the generation of a massive amount of training models to cover the different characteristics of the two models and the implementation of codes that can generate in a short time several thousand geological models with different structures. Ultimately, the deep learning technique, like other inversion methods, depends on the prior model.

The amount of data used in training also controls forecast accuracy, which generally improves as the size of training data increases. The use of a limited number of training models can lead to a poor generalization—this is known as the overfitting problem. Therefore, making reliable predictions requires very expensive numerical calculations and considerable computer resources to store large amounts of training data. Usually, the choice of the size of these data is based on a trial-and-error test until the required accuracy is achieved. This choice is dependent on the complexity of the training models and the ability of the apparent electrical resistivity measurements to cover these heterogeneities. In most cases, models with moderate heterogeneities are easy to train and predict, even with a modest amount of data, especially if the resistivity measurements provide high spatial resolution characterization. On the other

hand, when the electrical resistivity measurements are insufficient, noisy and unable to capture the heterogeneities of the field, it will be difficult to establish a relationship between the measurements and the model, leading to a failure of the learning and generalization operation. In this case of underfitting, increasing the size of training models will not improve the quality of forecasts.

In future work, we will focus on the application of this tool on the real field case with a comparative analysis between the results obtained with a classical deterministic inversion algorithm and our approach.

ACKNOWLEDGEMENTS

We would like to thank the editor and reviewers for their constructive comments and La région de Normandie for funding our research activities.

REFERENCES

- Apolinario, M.P.E., Huaman Bustamante, S.G., Morales, G. & Diaz, D., 2019. Estimation of 2D velocity model using acoustic signals and convolutional neural networks, in *IEEE XXVI International Conference on Electronics, Electrical Engineering and Computing (INTERCON)*, Lima, Peru, pp. 1–4.
- Araya-Polo, M., Dahlke, T., Frogner, C., Zhang, C., Poggio, T. & Hohlfeld, D., 2017. Automated fault detection without seismic processing, *Leading Edge*, **36**(3), 208–214.
- Araya-Polo, M., Jennings, J., Adler, A. & Dahlke, T., 2018. Deep-learning tomography, *Leading Edge*, **37**(1), 58–66.
- Badrinarayanan, V., Kendall, A. & Cipolla, R., 2017. SegNet: a deep convolutional encoder-decoder architecture for image segmentation, *IEEE Transactions on pattern analysis and machine intelligence*, **39**(12), 2481–2495, arXiv:1511.00561v3, <https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7803544>.
- Başokur, A.T. & Akca, I., 2011. Object-based model verification by a genetic algorithm approach: application in archeological targets, *J. Appl. Geophys.*, **74**(4), 167–174.
- Billi, A., De Filippis, L., Poncia, P.P., Sella, P. & Faccenna, C., 2016. Hidden sinkholes and karst cavities in the travertine plateau of a highly-populated geothermal seismic territory (Tivoli, central Italy), *Geomorphology*, **255**, 63–80.
- Cardarelli, E. & Di Filippo, G., 2009. Integrated geophysical methods for the characterisation of an archaeological site (Massenzio Basilica—Roman forum, Rome, Italy), *J. Appl. Geophys.*, **68**(4), 508–521.
- Cardarelli, E., Di Filippo, G. & Tuccinardi, E., 2006. Electrical resistivity tomography to detect buried cavities in Rome: a case study, *Near Surf. Geophys.*, **4**, 387–392.
- El-Qady, G. & Ushijima, K., 2001. Inversion of DC resistivity data using neural networks, *Geophys. Prospect.*, **49**(4), 417–430.
- Feibo, J., Li, D. & Qianwei, D., 2018. Electrical resistivity imaging inversion: an ISFLA trained kernel principal component wavelet neural network approach, *Neural Netw.*, **104**, 114–123.
- Fernández-Martínez, J.L., García-Gonzalo, E. & Naudet, V., 2010. Particle swarm optimization applied to solving and appraising the streaming-potential inverse problem, *Geophysics*, **75**(4), 3–15.
- Francisco, M.B., Walter, E.M. & Jerbeson, M.S., 2019. A user-driven feedback approach for 2D direct current resistivity inversion based on particle swarm optimization, *Geophysics*, **84**(2), 105–124.
- Furman, A., Ferre', T.P.A. & Warrick, A.W., 2004. Optimization of ERT surveys for monitoring transient hydrological events using perturbation sensitivity and genetic algorithms, *Vadose Zone J.*, **3**(4), 1230–1239.
- Günther, T., Rücker, C. & Spitzer, K., 2006. Three-dimensional modelling and inversion of dc resistivity data incorporating topography—II. Inversion, *Geophys. J. Int.*, **166**(2), 506–517.
- Hägström, I., Schmittlein, C.R., Campanella, G. & Fuchs, T.J., 2019. DeepPET: a deep encoder–decoder network for directly solving the PET image reconstruction inverse problem, *Med. Image Anal.*, **54**, 253–262.
- Hornik, K., Stinchcombe, M. & White, H., 1990. Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks, *Neural Netw.*, **3**(5), 551–560.
- Jardani, A., Revil, A. & Dupont, J.P., 2013. Stochastic joint inversion of hydrogeophysical data for salt tracer test monitoring and hydraulic conductivity imaging, *Adv. Water Res.*, **52**, 62–77.
- Kemna, A., Binley, A., Ramirez, A. & Daily, W., 2000. Complex resistivity tomography for environmental applications, *Chem. Eng. J.*, **7**, 11–18.
- Koefoed, O., 1979. *Geosounding Principles 1: Resistivity Measurements*, Elsevier Scientific Publication.
- LaBrecque, D.J. et al., 1996. ERT monitoring of environmental remediation processes, *Meas. Sci. Technol.*, **7**(3), 375–383.
- Lary, D.J., Alavi, A.H., Gandomi, A.H. & Walker, A.L., 2016. Machine learning in geosciences and remote sensing, *Geosci. Front.*, **7**(1), 3–10.
- LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P., 1998. Gradient-based learning applied to document recognition, *Proc. IEEE*, **86**(11), 2278–2324.
- Li, S., Liu, B., Ren, Y., Chen, Y., Yang, S., Wang, Y. & Jiang, P., 2020. Deep learning inversion of seismic data, *IEEE Trans. Geosci. Remote Sens.*, **58**, 2135–2149.
- Liu, B., Guo, Q., Li, S., Liu, B., Ren, Y., Pang, Y., Liu, L. & Jiang, P., 2020. Deep learning inversion of electrical resistivity data, *IEEE Trans. Geosci. Remote Sens.*, **58**, 5715–5728.
- Loke, M.H. & Barker, R.D., 1996. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method, *Geophys. Prospect.*, **44**(1), 131–152.
- Loke, M.H. et al., 2013. Recent developments in the direct-current geoelectrical imaging method, *J. Appl. Geophys.*, **95**, 135–156.
- Moseley, B., Nissen-Meyer, T. & Markham, A., 2020. Deep learning for fast simulation of seismic waves in complex media, *Solid Earth*, **11**, 1527–1549.
- Negri, S., Leucci, G. & Mazzone, F., 2008. High resolution 3D ERT to help GPR data interpretation for researching archaeological items in a geologically complex subsurface, *J. Appl. Geophys.*, **65**(3), 111–120.
- Oldenburg, D.W. & Li, Y., 1999. Estimating depth of investigation in dc resistivity and IP surveys, *Geophysics*, **64**(2), 403–416.
- Park, S.K. & Van, G.P., 1991. Inversion of pole–pole data for 3-D resistivity structure beneath arrays of electrodes, *Geophysics*, **56**(7), 951–960.
- Puzyrev, V., 2019. Deep learning electromagnetic inversion with convolutional neural networks, *Geophys. J. Int.*, **218**, 817–832.
- Roy, A. & Apparao, A., 1971. Depth of investigation in direct current methods, *Geophysics*, **36**(5), 943–959.
- Sharma, S.P., 2012. VFSARES—a very fast simulated annealing FORTRAN program for interpretation of 1-D DC resistivity sounding data from various electrode arrays, *Comput. Geosci.*, **42**, 177–188.
- Sharma, S.P. & Biswas, A., 2013. Interpretation of self-potential anomaly over 2D inclined structure using very fast simulated annealing global optimization—an insight about ambiguity, *Geophysics*, **78**, 3–15.
- Waxman, M. & Smits, L., 1968. Electrical conductivities in oil-bearing shaly sands, *Soc. Pet. Eng.*, **8**(2), 107–122.
- Wei, Z. & Chen, X., 2019. Deep-learning schemes for full-wave nonlinear inverse scattering problems, *IEEE Trans. Geosci. Remote Sens.*, **57**(4), 1849–1860.
- Werkema, D.D., Jr. et al., 2003. Investigating the geoelectrical response of hydrocarbon contamination undergoing biodegradation, *Geophys. Res. Lett.*, **30**(12), doi:10.1029/2003GL017346.
- Wu, Y. & Lin, Y., 2018. InversionNet: a real-time and accurate full waveform inversion with convolutional neural network, *J. Acoust. Soc. Am.*, **144**(3), 1683–1683.
- Yang, F. & Ma, J., 2019. Deep-learning inversion: a next generation seismic velocity-model building method, *Geophysics*, **84**(4), 583–599.
- Yao, W., Zeng, Z., Lian, C. & Tang, H., 2018. Pixel-wise regression using U-Net and its application on pansharpening, *Neurocomputing*, **312**, 364–371.
- Youssef, A.M., El-Kaliouby, H. & Zabramawi, Y.A., 2012. Sinkhole detection using electrical resistivity tomography in Saudi Arabia, *J. Geophys. Eng.*, **9**(6), 655–663.