Coupled hydrogeophysical inversion through Ensemble Smoother with Multiple Data Assimilation and Convolutional Neural Network for contaminant plume reconstruction

Received: / Accepted:

Camilla Fagandini[1], Valeria Todaro[1], Cláudia Escada[2], Leonardo Azevedo[2], J. Jaime Gómez-Hernández[3], Andrea Zanini[1]

1. Department of Engineering and Architecture, University of Parma, Parco Area delle Scienze 181/A, 43124, Parma, Italy.

2. CERENA/DER, Universidade de Lisboa, Instituto Superior Técnico, Lisbon, Portugal.

3. Institute for Water and Environmental Engineering, Universitat Politècnica de València, 46022 Valencia, Spain.

Corresponding author: Camilla Fagandini. Phone: +39 0521 905931; camilla.fagandini@unipr.it

ORCID: Camilla Fagandini: 0000-0003-2312-9096; Valeria Todaro: 0000-0002-9313-6999; Cláudia Escada: 0000-0002-3855-5947; Leonardo Azevedo: 0000-0002-0677-079X; Andrea Zanini 0000-0002-4022-707X; J. Jaime Gómez-Hernández: 0000-0002-0720-2196

Abstract

In the field of groundwater, accurate delineation of contaminant plumes is critical for designing effective remediation strategies. Typically, this identification poses a challenge as it involves solving an inverse problem with limited concentration data available. To improve the understanding of contaminant behavior within aquifers, hydrogeophysics emerges as a powerful tool by enabling the combination of non-invasive geophysical techniques (i.e., electrical resistivity tomography -

ERT) and hydrological variables. This paper investigates the potential of the Ensemble Smoother with Multiple Data Assimilation (ES-MDA) method to address the inverse problem at hand by simultaneously assimilating observed ERT data and scattered concentration values from monitoring wells. A novelty aspect is the integration of a Convolutional Neural Network (CNN) to replace and expedite the expensive geophysical forward model. The proposed approach is applied to a synthetic case study, simulating a tracer test in an unconfined aquifer. Five scenarios are compared, allowing to explore the effects of combining multiple data sources and their abundance. The outcomes highlight the efficacy of the proposed approach in estimating the spatial distribution of a concentration plume. Notably, the scenario integrating apparent resistivity with concentration values emerges as the most promising, as long as there are enough concentration data. This underlines the importance of adopting a comprehensive approach to tracer plume mapping by leveraging different types of information. Additionally, a comparison was conducted between the inverse procedure solved using the full geophysical forward model and the CNN model, showcasing comparable performance in terms of results, but with a significant acceleration in computational time.

Keywords: Inverse modeling; Ensemble Smoother; Groundwater contaminant source; Electrical Resistivity Tomography; CNN

1. Introduction

Over the last century, groundwater systems have faced increasingly severe environmental pressures as a consequence of massive industrial and agricultural development. To release these pressures, collaborative efforts are necessary, involving coordination with authorities and end-users to formulate decisions that prevent the depletion and contamination of aquifers. This poses a challenge that necessitates a comprehensive understanding of the subsurface environment and groundwater

systems whose complex spatial distribution can be difficult to characterize (Gómez-Hernández and Wen, 1994; Gómez-Hernández et al., 2003). Conventional survey methods, such as water sampling from monitoring wells, may not adequately capture a contaminant plume's structure and spread since they provide little localized information; furthermore, they are invasive, relatively expensive, and time-consuming. As a result, complementary techniques have been developed to overcome these survey-related challenges. Hydrogeophysics has emerged as a powerful, non-invasive and cost-effective tool in the field of contaminant hydrogeology, leveraging geophysical data to gain insights into hydrological processes and the underlying geology that govern the subsurface (Rubin and Hubbard 2005; Vereecken et al. 2006; Hubbard and Linde 2011). These methods, such as electrical resistivity tomography (ERT), ground-penetrating radar (GPR), and seismic surveys, enable subsurface imaging and detection of anomalies. Given that polluted groundwater exhibits increased electrical conductivity (Frohlich and Urish 2002; Carpenter et al. 2012), approaches that measure ground electrical conductivity or its reciprocal, electrical resistivity, become particularly interesting when combined with hydrological data. For this reason, ERT is widely used in hydrological studies (e.g., Page 1968; Wilson et al. 2006; Pereira et al. 2023).

Recovering aquifer characteristics and groundwater contaminant information from geophysical data, alongside sparse hydrological data, requires solving a complex geophysical inverse problem. Several deterministic and stochastic methods have been proposed to address these challenges. A comprehensive review of hydrogeology inverse methodologies is available in the works of McLaughlin and Townley (1996), Zimmerman et al. (1998), Carrera et al. (2005), Hendricks Franssen et al. (2009), Zhou et al. (2014) and Gómez-Hernández and Xu (2022). Stochastic inverse methods, such as the geostatistical approach (Kitanidis 1995), offer an effective way of characterizing spatial variability and inferring properties of interest at unsampled locations associated with their uncertainty (Michalak and Kitanidis 2004; El Idrysy and De Smedt 2007; Huysmans and Dassargues 2009; Zhou et al., 2012; Butera et al. 2013; Cupola et al. 2015; Zanini and Woodbury 2016; Visentini et al. 2020). Among the stochastic inversion techniques, the

ensemble Kalman filter (Evensen 1994) and the ensemble smoother (Leeuwen and Evensen, 1996), have seen a rise in popularity in hydrogeology due to their adaptability and effectiveness (Chen and Zhang 2006; Li et al. 2012; Crestani et al. 2013, 2015; Xu and Gómez-Hernández, 2016, 2018; Chen et al. 2018, Li et al. 2019). In particular, Emerick and Reynolds (2012, 2013) introduced the ensemble smoother with multiple data assimilation (ES-MDA), which involves the iterative assimilation of the same data multiple times, enhancing the applicability and efficacy of the ensemble smoother (Todaro et al. 2019, 2021, 2023; D'Oria et. al, 2021; Xu et al., 2021; Godoy et al. 2022; Chen et al. 2023).

Several works have shown how hydrogeophysics inverse modeling can be used in conjunction with ERT measurements to estimate hydraulic properties such as hydraulic conductivity (Irving et al., 2010; Pollock and Cirpka, 2010, 2012), including the use of Kalman-based techniques (Kang et al., 2019; Camporese et al., 2011, 2015; Crestani et al., 2015). However, few studies have focused on utilizing ERT measurements to predict groundwater contamination. Kang et al. (2018) employed the ensemble Kalman filter to simultaneously estimate the distribution of dense non-aqueous phase liquid (DNAPL) saturation and aquifer heterogeneous parameter field using time-lapse ERT data. Tso et al. (2020) employed ES-MDA to detect contaminant leaks utilizing time-lapse ERT measurements. Chen et al. (2023) utilized the ES-MDA to jointly identify contaminant source information and the hydraulic conductivity field by assimilating ERT data in a synthetic heterogeneous aquifer with a time-varying release history. The results underscored the capability of the ES-MDA data assimilation framework to provide a robust inversion of both time-varying release history and hydraulic conductivity estimation.

The aforementioned research findings demonstrated hydrogeophysics' ability to identify pollutant sources and aquifer characteristics. However, one major challenge in inverse modeling is the complexity of the underlying forward models, which are often computationally expensive or analytically unsolvable. Surrogate models present a viable solution to overcome these issues (e.g., Asher et al. 2015; Jamshidi et al., 2020; Secci et al., 2022, 2024). In recent years, neural networks

have emerged as a promising tool for replacing full forward models and reducing computational demand. A well-known neural network is the convolutional neural network (CNN) introduced by LeCun et al. (1998). CNNs specialize in processing grid-based data, such as images, exhibiting an inherent capacity to capture and hierarchically represent spatial features in data. For this reason, CNN is a technology widely employed in various fields, including groundwater spatial modeling (e.g., Hong and Liu, 2020; Panahi et al., 2020; Lähivaara et al., 2019).

In the literature, only a few studies have explored the potential of coupling CNN with ES-MDA. Tang et al. (2021) combined convolutional post-processing of principal component analysis parameterization and ES-MDA to estimate both a channelized permeability and oil/water rate in petroleum engineering. Zhou et al. (2022) integrated convolutional adversarial autoencoder and ES-MDA to parameterize a non-Gaussian conductivity field and to identify the spatiotemporal extended source of contamination. In this work, the ES-MDA and CNN are coupled to unlock the potential of hydrogeophysics in addressing environmental pollution problems while lowering the computational cost of the inversion procedure. The primary objective is to combine hydrological and ERT data to accurately estimate the spatial distribution of a contaminant within a groundwater system.

The ES-MDA inverse procedure is applied to estimate the plume distribution by employing a well-established geophysical forward model and assimilating both ERT data and sparse concentration values from monitoring points. To enhance efficiency, a CNN is used to replace the part of the forward model that transforms the electrical resistivity of the investigated material into the apparent electrical resistivity that would be deduced from an ERT survey. The proposed methodology is tested by means of a two-dimensional synthetic case study that mimics a tracer test in an unconfined aquifer. Different scenarios are investigated exploring the effect of combining multiple data sources and their abundance.

The structure of this paper is outlined as follows. Section 2 provides a comprehensive overview of the material and methods employed in the proposed inverse approach. Section 3 details the test case set up, the configuration of the CNN end the ES-MDA, as well as the investigated scenarios.

Section 4 delves into the presentation and analysis of results. Finally, Section 5 presents discussions and conclusions.

2. Material and Methods

2.1 Forward model

The forward model has two components. The first is a petrophysical model used to spatially predict the resistivity field associated with a given contaminant plume. The second is a geophysical model utilized to calculate the apparent resistivity (i.e., pseudo-electrical resistivity) that would be observed during an Electrical Resistivity Tomography (ERT) survey associated with a given subsurface electrical resistivity field. In this work, the geophysical model is replaced by a convolutional neural network. The following sections describe the entire forward model in detail.

2.1.1 Petrophysical relationship

Petrophysical models are needed to link geophysical imaging techniques and hydrological models (Vereecken et al. 2006). In this case, the model proposed by Pollock and Cirpka (2012) is used to transform concentration into electrical conductivity (EC) using

$$\sigma(t,\mathbf{x}) = \sigma_0(\mathbf{x}) + \sigma'(t,\mathbf{x}) \quad (1)$$

where $\sigma(t,\mathbf{x})$ is the bulk electrical conductivity at specific time t and location $\mathbf{x}$, $\sigma_0(\mathbf{x})$ is the background bulk electrical conductivity (constant through time), and $\sigma'(t,\mathbf{x})$ is a perturbation resulting from a change in solute concentration $c(t,\mathbf{x})$. $\sigma'(t,\mathbf{x})$ can be derived from Archie's law (Archie 1942)

$$\sigma'(t,\mathbf{x}) = \frac{\varphi^m}{a} S^n \sigma_w \, c(t,\mathbf{x}) \quad (2)$$

with φ being the porosity, S being the water saturation, σ_w being the water EC, m and n being two empirical parameters referred to as cementation and saturation exponent, respectively. a is a proportionality constant of the order of 1.

153 Electrical resistivity (ρ) is the reciprocal of EC

154 $$\rho = \frac{1}{\sigma} \quad (3)$$

155

156 **2.1.2 Electrical Resistivity Tomography (ERT): governing equations**

157 A common ERT survey considers four electrodes and consists of injecting electrical current into the
158 ground through two current electrodes (C1 and C2) and measuring the resulting voltage difference
159 at two potential electrodes (P1 and P2). Afterward, the current and voltage measurements are
160 transformed into apparent electrical resistivity, which represents a weighted average resistance of
161 earth materials to electrical current propagation (Loke et al. 2013).

162 Poisson's equation can be used to describe the electric potential field generated by a couple of
163 electrodes

164 $$-\nabla \cdot \sigma(x,y,z)\nabla\phi(x,y,z) = I\big(\delta(r - r_+) - \delta(r - r_-)\big) \quad (4)$$

165 in which ϕ represents the potential field; I is the input current; r_+ and r_- are the locations of the
166 positive and negative electrodes, respectively, and $\delta(\cdot)$ is the Dirac delta function. Following
167 Pidliskey and Knight (2008), the solution to Eq. 4 yields a vector of electric potential values for
168 each grid location within the considered model. Then, for a given electrode array, the apparent
169 electrical resistivity at a location in the xz plane that is specific to such configuration is computed as

170 $$\rho_{app} = \Delta\widehat{\phi} \cdot K \quad (5)$$

171 where $\Delta\widehat{\phi}$ is the difference of potential recorded between the electrodes P1 and P2, and K is a
172 geometric factor, which is a function of the distance among the four electrodes calculated as follows
173 when the effects of the topography are ignored

174 $$K = \frac{2\pi}{\frac{1}{d_1} - \frac{1}{d_2} - \frac{1}{d_3} + \frac{1}{d_4}} \quad (6)$$

175 where d_1 is the distance between the current electrode C1 and the potential electrode P1, d_2 is the
176 distance between the current electrode C1 and the potential electrode P2, d_3 is the distance between

177 the current electrode C2 and the potential electrode P1, d_4 is the distance between the current
178 electrode C2 and the potential electrode P2.
179 The apparent electrical resistivity values are then visualized in a 2D "pseudo-section" plot,
180 providing a comprehensive view of both horizontal and vertical changes. The horizontal position of
181 each data point corresponds to the midpoint of the electrode set used for measurement, while its
182 vertical position represents a proportionate distance based on electrode separation. For further
183 insight into the specific array configuration, readers are directed to Edwards (1977).
184 According to Pidliskey and Knight (2008) and assuming no variation along the y-axis
185 $(\frac{\partial}{\partial y}\sigma(x, y, z) = 0)$, a 2.5D forward ERT model, is used to calculate the apparent electrical resistivity
186 from an electrical resistivity model. The forward geophysical problem is solved using SimPEG
187 (Cockett et al., 2015), an open-source geophysical library.
188
189 **2.1.3 Surrogate model: Convolutional Neural Network (CNN)**
190 Convolutional Neural Networks (CNNs), first developed by LeCun et al. (1998), represent a class
191 of machine learning models designed for processing and analyzing visual data, making them
192 particularly effective for tasks involving images or spatially structured data. At their core, CNNs
193 leverage convolutional filters: small learnable matrices that slide over the input image, capturing
194 spatial hierarchies and local patterns. This allows CNNs to efficiently recognize complex patterns
195 and spatial relationships within the data. Several review papers have been presented in the last few
196 years, offering comprehensive overviews of the CNN advancements and applications (see e.g., Gu
197 et al. 2018; Khan et al, 2020; Alzubaidi et al., 2021). Within the geophysical inversion context,
198 CCNs have been utilized in studies such as Das et al. (2019) and Puzyrev (2019). A CNN comprises
199 an input layer, several hidden layers, and an output layer. The input layer receives the raw input
200 data in the form of images or other grid-like data. Typically, CNN hidden layers consist of
201 convolutional layers, activation functions, pooling layers, and possibly batch normalization.

Convolutional layers apply filters to capture local features. The use of activation functions, such as rectified linear units (ReLU), introduces non-linearity to the model, enhancing its ability to capture intricate patterns. Pooling layers with specified pool sizes and strides downsample the spatial dimensions, reducing computational complexity. Batch normalization may be included to normalize the input activations, enhancing training stability. The CNN architecture typically concludes with a fully connected (dense) layer, which takes the features learned by the convolutional layers and combines them to make predictions. Dropout layers can also be included to mitigate overfitting by randomly deactivating a fraction of neurons during training. Ultimately, the output layer produces the final prediction. The training process involves iteratively adjusting the weights of the network using optimization algorithms, such as Adam optimizer (Zhang 2018), to minimize the difference between predicted and target values.

In this study, a CNN is employed to replace the electrical resistivity forward model described in the previous section. The input layer comprises a resistivity map, and the output layer yields apparent resistivity data. The details of the CNN's architecture employed for this particular application are outlined in the Section 3.3.

2.2 ES-MDA inversion approach

The method applied to solve the hydrogeophysical inverse problem is the ensemble smoother with multiple data assimilation (ES-MDA). The ES-MDA is an iterative data assimilation approach that allows the estimation of model parameters using a set of observed measurements and a known relationship between parameters and observations, given by a forward model. A brief description of the method is provided next. For a more detailed description, the reader is referred to Emerick and Reynolds (2013).

The method workflow consists of an initialization phase and an iterative phase; in which each iteration is made up of two steps: a forecast step and an update step. The initialization phase involves the generation of an initial ensemble of parameter realizations $\mathbf{X} \in \Re^{Np \times Ne}$, where N_p is the

number of parameters to be estimated and N_e is the ensemble size, together with an ensemble of observation errors $\boldsymbol{\varepsilon} \in \Re^{m \times Ne}$, where m is the number of observations. Moreover, the procedure requires the definition of a priori number of iterations N and a vector of inflation coefficients $\{\alpha_i, i = 1,...,N\}$. Several schemes can be used to define the set of α, but they must satisfy the condition

$$\sum_{i=1}^{N} \frac{1}{\alpha_i} = 1 \tag{7}$$

After the initialization step, iterations start. During the forecast step, at each iteration i, for each realization j of the ensemble of parameters $\mathbf{X}_{j,i} \in \Re^{Np}$, the forward model is run to obtain the model predictions, of which a subset $\mathbf{Y}_{j,i} \in \Re^{m}$ is extracted coinciding with the same locations and times as the observations $\mathbf{D} \in \Re^{m}$

$$\mathbf{Y}_{j,i} = g(\mathbf{X}_{j,i}) \tag{8}$$

where g (·) is an operator that incorporates the forward model as well as a filtering function used to extract the predictions at the m locations where observations have been collected. Next, the ensemble of parameters is updated during the update step according to the equation

$$\mathbf{X}_{j,i+1} = \mathbf{X}_{j,i} + \mathbf{Q}_{\mathbf{XY}}^{i}\left(\mathbf{Q}_{\mathbf{YY}}^{i} + \alpha_i \mathbf{R}\right)^{-1}\left(\mathbf{D} + \sqrt{\alpha_i}\boldsymbol{\varepsilon}_j - \mathbf{Y}_{j,i}\right) \tag{9}$$

where $\mathbf{Q}_{\mathbf{XY}}^{i} \in \Re^{Np \times m}$ is the cross-covariance matrix between parameters and predictions, $\mathbf{Q}_{\mathbf{YY}}^{i} \in \Re^{m \times m}$ is the auto-covariance matrix of predictions and $\mathbf{R} \in \Re^{m \times m}$ is the auto-covariance matrix of the measurement errors, which are assumed to be uncorrelated. $\boldsymbol{\varepsilon}_j \in \Re^{m}$ is the vector of measurement errors for realization j. $\mathbf{Q}_{\mathbf{XY}}^{i}$ and $\mathbf{Q}_{\mathbf{YY}}^{i}$ are computed, from the ensemble of realizations, at each iteration i as

$$\mathbf{Q}_{\mathbf{XY}}^{i} = \frac{1}{N_e - 1}\sum_{j=1}^{N_e}\left(\mathbf{X}_{j,i} - \overline{\mathbf{X}}_i\right)\left(\mathbf{Y}_{j,i} - \overline{\mathbf{Y}}_i\right)^{T} \tag{10}$$

$$\mathbf{Q}_{\mathbf{YY}}^{i} = \frac{1}{N_e - 1}\sum_{j=1}^{N_e}\left(\mathbf{Y}_{j,i} - \overline{\mathbf{Y}}_i\right)\left(\mathbf{Y}_{j,i} - \overline{\mathbf{Y}}_i\right)^{T} \tag{11}$$

where $\overline{\mathbf{X}}_i$ and $\overline{\mathbf{Y}}_i$ are the ensemble means, at iteration *i*, of **X** and **Y**, respectively.

The iteration index then advances, and the algorithm returns to the forecast step until the final iteration.

252 To minimize the number of parameter realizations, since computation time depends on it,
253 covariance and inflation techniques are employed. These methods help prevent issues stemming
254 from small ensemble sizes. The covariance localization involves an element-wise multiplication of
255 the original covariance matrices with selected tapering functions that diminish correlations between
256 points as the distance increases, effectively eliminating spurious long-range spurious correlations
257 beyond a specified threshold. Covariance inflation is additionally taken into account to address
258 issues related to under sampling. At each iteration, it modifies the ensemble of updated parameters
259 by adjusting the ensemble spread, preventing the spread from becoming too narrow with the
260 consequence of collapse and divergence.
261 The software package genES-MDA developed by Todaro et al. (2022) is used to apply the ES-
262 MDA procedure.
263

264 **2.3 Coupled hydrogeophysical inverse model**

265 This section summarizes the scheme of the proposed coupled hydrogeophysical inversion, which
266 seeks to estimate the spatial distribution of a tracer plume by integrating available observations (e.g.
267 observed ERT data and concentration values at monitoring points). The methodology comprises
268 several steps detailed below (Fig. 1).

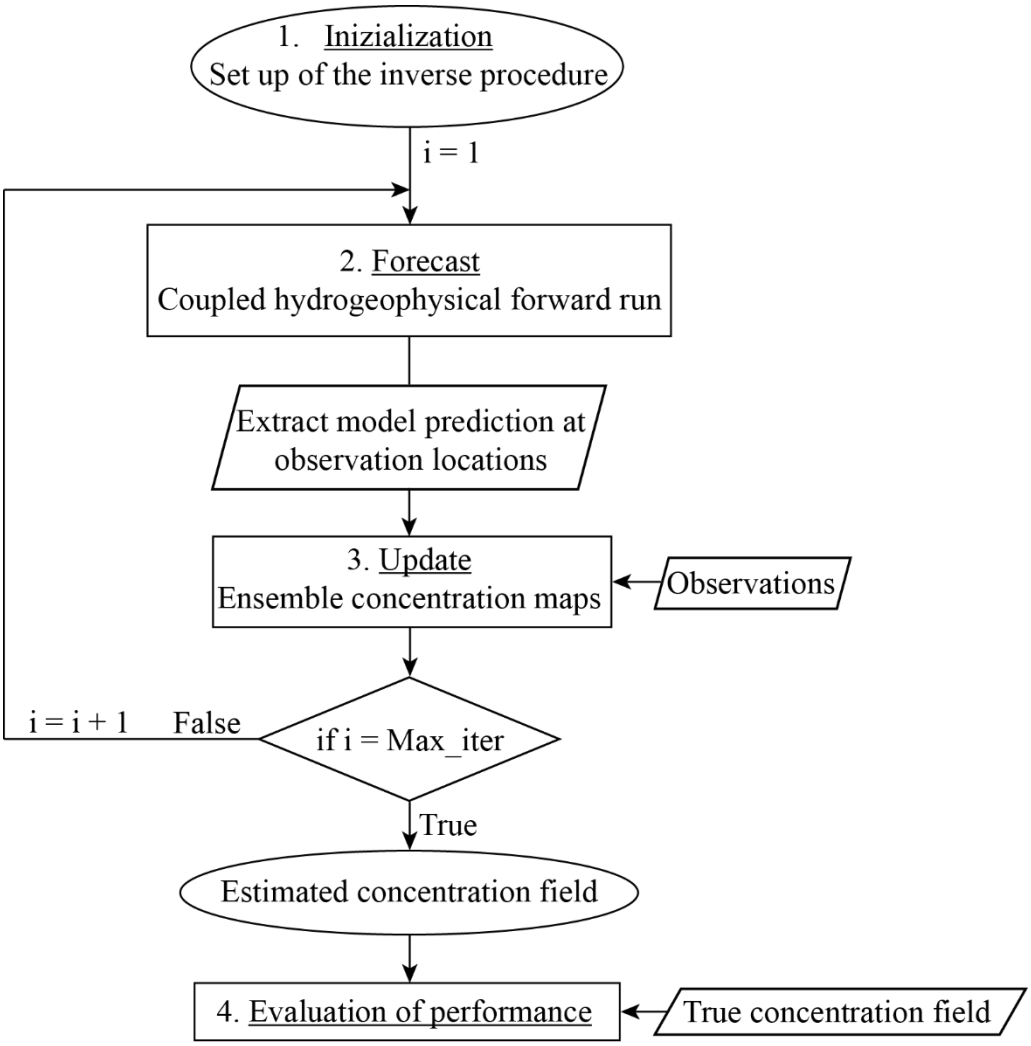


Fig. 1 Flowchart of the coupled hydrogeophysical inverse model

Step 1. Initialization

The first step involves the generation of the initial parameter realizations. These realizations correspond to different concentration fields, aiming to incorporate available a priori information and adequately represent the specific problem under consideration. The initial concentration maps can be systematically generated through various approaches, ensuring a comprehensive exploration of the subsurface conditions, one can:

i) Assume homogeneity across all parameters. In this scenario, each realization exhibits a distinct constant value drawn from a uniform distribution. This method is straightforward and feasible in situations where no prior information is available.

ii) Run stochastic sequential simulations to generate fields using a semi-variogram model. The semi-variogram can be fitted to existing field data if available, or alternatively, variogram parameters can be selected randomly from a range defined based on prior knowledge. This approach considers the spatial correlations present in the reference dataset, ensuring that the initial ensemble captures the expected patterns of the actual concentration map.

iii) Utilize a numerical transport model to generate diverse realizations by simulating the injection from different locations within the model domain as well as various tracer concentrations, both randomly selected from predefined tailored ranges. This ensures that each realisation considers realistic representation of contaminant distribution in the subsurface.

This step also involves the definition of the number of iterations N, the observation errors, the coefficients α_i and the training of the CNN.

Step 2. Forecast: CNN-based forward model

For each iteration and for every realization, the petrophysical relationship, described in Section 2.1.1 is used to transform the concentration maps into electrical resistivity maps. Subsequently, these maps undergo forward modelling with the trained CNN, resulting in apparent electrical resistivity values. A filtering function is employed to extract the subset of prediction data at the observation locations.

Step 3. Update

At each iteration, the prediction vector is used to update the concentration map following Eqs. 9-11. Upon completing the concentration update, the subsequent iteration starts with the updated ensemble of parameters. Step 2 and Step 3 are repeated until the last iteration.

Step 4. Analysis and interpretation of the results

The results are analyzed in terms of mean and standard deviation computed from the ensemble, allowing to associate the parameter estimation with their uncertainty.

If a reference solution is available, as is common in synthetic case studies, a thorough comparison is made between the estimated and actual concentration values. The assessment of results employs well-established metrics, specifically, the mean error (ME), the mean absolute error (MAE), the root mean squared error (RMSE) and the determination coefficient (R^2) as given by

$$ME = \frac{1}{N_p}\sum_{k=1}^{N_p}\left(\hat{C}_k - C_k\right) \quad (12)$$

$$MAE = \frac{1}{N_p}\sum_{k=1}^{N_p}\left|\hat{C}_k - C_k\right| \quad (13)$$

$$RMSE = \sqrt{\frac{1}{N_p}\sum_{k=1}^{N_p}\left(\hat{C}_k - C_k\right)^2} \quad (14)$$

$$R^2 = 1 - \frac{\sum_{k=1}^{N_p}\left(\hat{C}_k - C_k\right)^2}{\sum_{i=1}^{N_p}(C_k - \overline{C_k})^2} \quad (15)$$

where N_p is the number of parameters (in this case is the number of grid nodes), C_k is the actual concentration, $\hat{C}_k$ is the estimated ensemble-mean concentration and $\overline{C_k}$ is mean actual concentration.

3. Application

3.1 Set up of the Test Case

The validity of the proposed methodology is demonstrated by its application to a two-dimensional synthetic model representing the vertical cross section of a heterogeneous unconfined aquifer under fully saturated conditions, where a contaminant plume is present. This model resembles the sandbox developed at the University of Parma's Hydraulic Laboratory, which has been extensively used in experimental and computational studies (Citarella et al. 2015; Cupola et al. 2015; Chen et al. 2018, 2021; Todaro et al. 2021, 2023; Pereira et al. 2023).

Fig. 2 offers a schematic depiction of the synthetic model being discussed. It is discretized into 96 by 1 by 20 cells, each measuring 1 by 10 by 1 cm. The hydraulic conductivity varies in space with three well-defined homogeneous zones differing by two orders of magnitude (Figure 2 and Table

1), and a uniform porosity equal to 0.37. The boundary conditions are impermeable at the bottom, phreatic surface at the top, and fixed heads at the left and right sides. This setup generates a head loss of 1 cm that induces flow from left to right. The initial condition is zero concentration everywhere. A continuous injection of a conservative non-reactive tracer, with a concentration of 20 mg/L, is introduced into the model from a designated injector point at location (X=12.5 cm, Z =10.5 cm). Longitudinal and transverse dispersivity values are assumed to be 0.16 cm and 0.016 cm, respectively. The reference solution is derived from a simulation conducted using MODFLOW (Harbaugh 2005) and MT3DMS (Zheng and Wang 1999) to model the groundwater flow and mass transport process, respectively. Table 1 summarizes the model parameters. The simulation extends for a total duration of 3600 seconds to achieve a well-developed plume, with the concentration map at the final time step serving as the reference map. The parameters to be estimated correspond to the concentration at each model grid cell (N_p=1920).

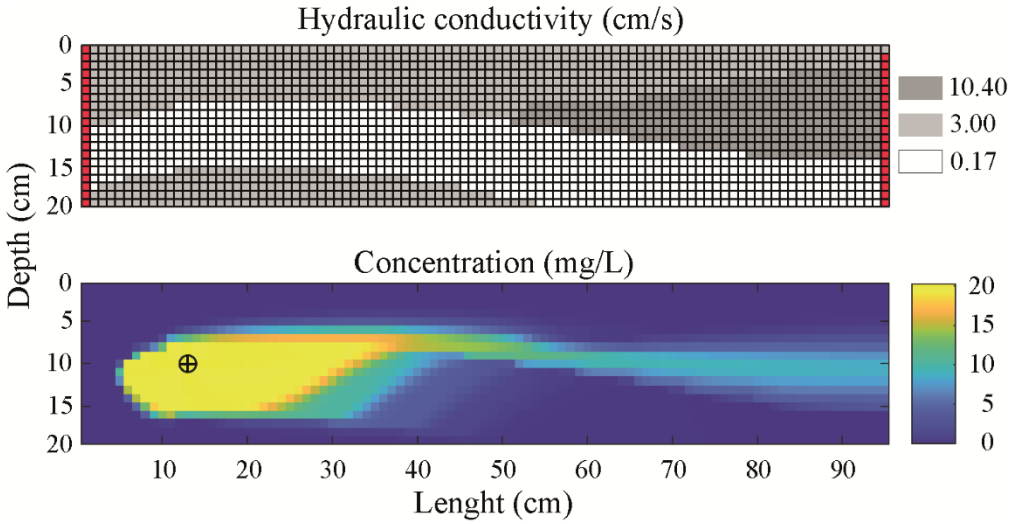


Fig. 2 Hydraulic conductivity and concentration reference maps. The red grid cells represent the left and right boundary conditions. The cross indicates the injector position

Table 1 Flow and transport model parameters

Hydraulic Conductivity 1 (cm/s)	0.17
Hydraulic Conductivity 2 (cm/s)	3.00
Hydraulic Conductivity 3 (cm/s)	10.40
Porosity	0.37

Longitudinal dispersivity (cm)	0.16
Transverse dispersivity (cm)	0.016
Injected concentration (mg/L)	20

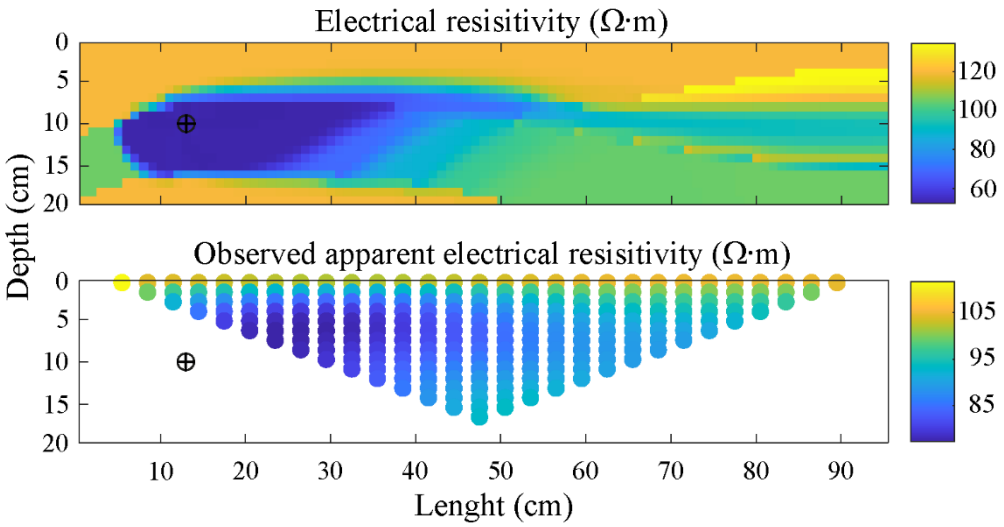
347

348 The reference electrical resistivity map (Fig. 3) is obtained by applying the petrophysical model
349 described in Section 2.1.2. Then, the SimPEG package processes the resulting map to derive the
350 apparent resistivity at 225 locations, representing the observations to be used in the inverse
351 procedure. This estimate is made using Eqs 4-6 and taking into account a Wenner-Schlumberger
352 acquisition array, which consists of 32 electrodes spaced at 3 cm intervals.
353 Table 2 summarizes the geophysical and petrophysical parameters used.

354 **Table 2** Geophysical and petrophysical parameters

Number of electrodes	32
Electrode spacing (cm)	3
m	1.3*
n	2
a	1
S_w	1
σ_w (µS/cm)	357

355 * (Mavko et al. 2009)



356

357 **Fig. 3** Reference resistivity model and observed pseudo-section. The cross indicates the position of
358 the injector

359

360 **3.2 Investigated scenarios**

The idea of the work rises from the necessity to visualize the plume spread into aquifers. One possibility is the interpolation of observed concentrations in the field if they are available. Normally these data are few and spatially sparse. Therefore, the introduction of ERT measurements, which are spatially exhaustive, is considered. In order to assess the capabilities of the proposed approach, five distinct scenarios considering different datasets are developed. Each dataset aimed to emphasize the advantages of employing specific combinations of apparent resistivity measurements (m_1) and concentration measurements (m_2).

Three monitoring wells are placed along the vertical at x = 23.5, 47.5, and 71.5 cm, each with five equidistant observation points spaced at 3 cm interval, for a total of 15 observation points. In Scenario 1, a limited dataset, comprising only the 15 concentration values, is used to interpolate the concentration map. This map is generated using a kriging-based interpolation method, with the variogram model computed using the 15 concentration observations ($m_1 = 0$, m_2=15). The intent is to demonstrate how difficult is to obtain a good estimate using a spatially sparse data set. In the other scenarios, parameters are estimated in each cell of the model grid using the ES-MDA hydrogeophysical inversion, with the number of observations varying according to the specific case under examination. In Scenario 2 ($m_1 = 225$, $m_2 = 0$) only ERT data are employed. In Scenario 3 ($m_1 = 225$, $m_2 = 15$) the ERT data are combined with 15 concentration values. In Scenario 4 ($m_1 = 225$, $m_2 = 9$) the ERT data are combined with 9 concentration values. And, in Scenario 5 ($m_1 = 225$, $m_2 = 3$), the ERT data are combined with only 3 concentration values. A summary of the five scenarios is provided in Table 3.

Table 3 Summary of the scenarios. Number of observations used.

Scenario	1	2	3	4	5
ERT data	-	225	225	225	225
Concentration data	15	-	15	9	3

3.3 CNN's set up

To speed up the execution of the forward model, a CNN is implemented to replace the SimPEG package that converts electrical resistivity into apparent electrical resistivity data (i.e., pseudo-electrical resistivity sections). To train the network, a dataset including 7000 realizations obtained with SimPEG, is considered. This input dataset undergoes preprocessing involving the normalization of input and output data and is then split into training (70%), validation (15%), and test (15%) sets. The CNN architecture is outlined in Table 4. The model is optimized using the Adam optimizer with a learning rate of 0.001, and the mean squared error between predicted and target apparent resistivity values is used as the loss function. The training is performed with a batch size of 18 over 300 epochs. After training, the model is evaluated on the validation set, and predictions are inverse transformed to the original scale. The complete CNN training and validation process tooks approximately 3 hours utilizing a computer equipped with Intel i9-10920X CPU 3.5GHz, 32 GB RAM.

Table 4 CNN architecture

Layer	Number of filters	Size of each filter	Stride	Padding	Batch Normalization	Activation	Output size
Input image							20x96x1
Convolutional	8	5x5x1	1x1	Same	True	ReLU	20x96x8
Pooling (Average)	-	2x2	2x2	0	False	None	10x48x8
Convolutional	16	5x5x8	1x1	Same	True	ReLU	10x48x16
Pooling (Average)	-	2x2	2x2	0	False	None	5x24x16
Convolutional	32	5x5x16	1x1	Same	True	ReLU	5x24x32
Convolutional	64	5x5x32	1x1	Same	True	ReLU	5x24x64
Dropout (50%)	-	-	-	-	False	None	5x24x64
Fully connected	-	-	-	-	False	Linear	1x225

Fig. 4 reports the results of the validation set. It is clear the good agreement between the true and computed apparent resistivities.

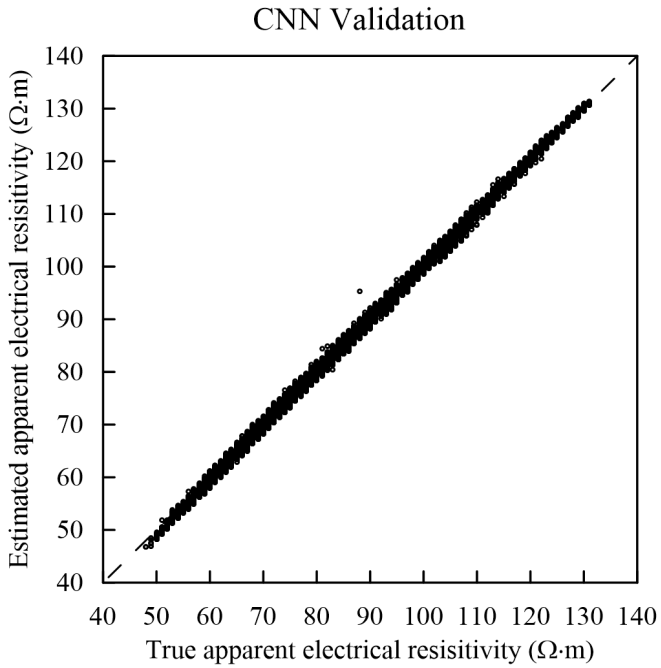


Fig. 4 CNN Validation, the dashed line is the 1:1 line

The computational time of the CNN was compared to that of the 2.5D forward ERT model, showing a substantial reduction for each realization from 2.3 seconds to approximately 20 milliseconds using the computational infrastructure described above.

Furthermore, to validate the reliability of the presented approach, the inverse problem in Scenario 3 is solved by using both the CNN model and the full forward model (SimPEG), comparing their performance.

3.4 Inverse model set up

For Scenarios 2-5, the ES-MA is performed with six iterations and an ensemble size of 500. In this study, the initial ensemble of parameters is generated following approach ii) described in Section 2.3 (Step 1) using the Python package GeostatsPy (Pyrcz et al. 2021), which interfaces the Geostatistical Software Library (GSLIB) with Python. It is employed to generate Gaussian random fields in logarithmic space to prevent negative values, which are subsequently back-transformed into the concentration space. Each realization is based on an anisotropic exponential variogram, with an azimuth for the largest continuity set at 90 degrees. The mean log-concentration is randomly selected

418 from a uniform distribution within the interval [-2, 2], while the standard deviation is equal to 1.1.
419 The correlation range in the vertical direction is randomly selected from a uniform distribution with
420 ranges [10, 20] (cm), while the anisotropy ratio is sampled within the range [7, 10] (cm).
421 The observation errors are normally distributed with zero mean and variance equal to 10^{-4} $(\Omega\ m)^2$ for
422 the apparent electrical resistivity and 0.01 $(mg/L)^2$ for the concentrations. A decreasing α set with
423 values equal to [364.0; 121.3; 40.4; 13.5; 4.5; 1.5] is used. A spatial covariance localization is applied
424 considering a space correlation length of 30 cm. A covariance inflation is applied with a factor equal
425 to 1.01 (refer to Todaro et al., (2023) for a detailed explanation of ES-MDA set up).

426 ## 4. Results

427 The comparative analysis of the five scenarios reveals different insights into the efficacy of ES-MDA
428 in estimating the distribution and values of the concentration plume for a given release history. The
429 results are depicted in Fig. 5, where the estimated concentration for Scenarios 2-5 is given by the
430 ensemble mean. Table 5 provides the evaluation metrics, assessed using Eqs. 12-15, alongside the
431 maximum estimated concentration for comparison with the actual value of 20 mg/L. Additionally, it
432 encompasses an assessment of estimate uncertainty as indicated by the standard deviation.
433 In the first scenario (Fig. 5.a), the concentration map is obtained through kriging interpolation using
434 15 concentration values; this result provides a baseline for performance evaluation. Moving on to
435 Scenario 2 (Fig. 5.b), where ES-MDA is utilized with only apparent resistivity as observations, the
436 results exhibit poorer accuracy in the estimation of the concentration map, compared to the previous
437 one (Fig. 2). While the estimation of the contaminant distribution is satisfactory and the RMSE of
438 3.69 mg/L is comparable to that of Scenario 1, there is a significant overestimation of the injected
439 concentration, resulting in higher mean error (-0.48 mg/L) and mean absolute error (2.64 mg/L). In
440 particular, the maximum estimated concentration reaches around 32 mg/L, whereas the actual
441 concentration is 20 mg/L. The absence of concentration data highlights the significance of
442 incorporating such information for a more robust estimation. In comparison, the third scenario (Fig.

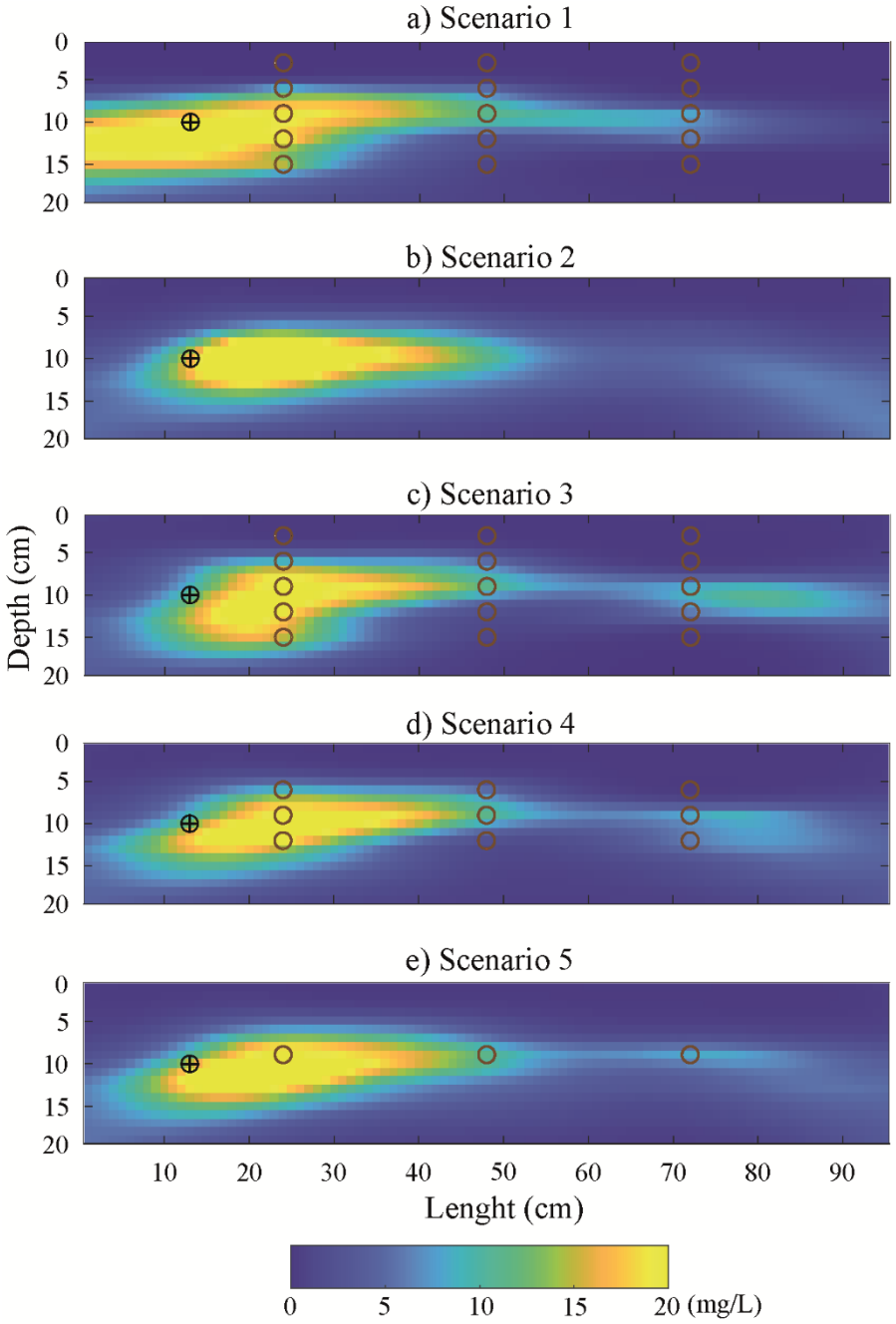
443 5.c), which combines apparent resistivity data and the 15 concentration values, emerges as the best
444 result in terms of observation estimation and field distribution. When compared to the other scenarios,
445 this integrated approach outperforms the previous ones with a ME = 0.06 mg/L; MAE = 1.56 mg/L;
446 RMSE = 2.74 mg/L; R^2 = 0.82, and the best estimate of the maximum concentration of 22 mg/L,
447 which is close to the actual injected. The combination of geophysical data and concentration values
448 improves the model's ability to capture plume distribution. Scenario 4 (Fig. 5.d), which is similar to
449 Scenario 3 but considers only 9 concentration observations, reveals a subtle trade-off between data
450 quantity and model accuracy. Although the reduction in concentration data slightly affects accuracy,
451 the overall performance remains good (ME = -0.01 mg/L; MAE = 1.97 mg/L; RMSE = 3.01 mg/L;
452 R^2 = 0.79). The limitations of the sparse concentration information become more pronounced in the
453 final scenario (Fig. 5.e), where only three concentration data points are used in conjunction with
454 apparent resistivity data. Despite the model's adaptability, the reduced data set compromises the
455 accuracy of the estimated concentration map, as indicated by ME of -0.38 mg/L, MAE of 2.21 mg/L,
456 RMSE of 3.17 mg/L, Max Concentration of 28.62 mg/L, and R^2 of 0.76.
457 Fig. 6 shows the scatterplot between true and estimated concentrations at each model grid cell
458 (N_p=1920) for all investigated scenarios. The dispersion data points indicates that there is not a
459 perfect agreement between the estimated and true values. Despite this dispersion, the best linear fit,
460 illustrated by the red line in Fig.6, indicates that the model's overall predictive ability is good, with
461 slopes ranging from 0.71 (Scenario 4) to 0.80 (Scenario 1). This is also supported by a high R^2 value
462 (Table 5). The results of the inversion procedure effectively capture a significant portion of the
463 variation in the true concentration field. However, as highlighted in Fig. 6, the methodology
464 encounters difficulties, particularly in identifying the lower and higher concentration values in some
465 scenarios, pointing out the limits of each application.. The interpolation in Scenario 1 faces a
466 challenge in accurately estimating lower values, while maximum values are quite well represented.
467 Comparing the true contaminant distribution (Fig. 2) and the estimated one (Fig.5a) it can be noticed
468 that the concentrations in the area upstream of the source location are overestimated, mainly due to

469 the extrapolation by the kriging estimator beyond the position of the available data. In Scenarios 2 to
470 5 the proposed procedure better estimates the lower values whereas it presents large uncertainty on
471 the maximum concentration (see Fig. 6). In particular, comparing the true concentration map with the
472 estimated one in Scenarios 2 to 5 (Fig. 5b-e), it is evident that most of the underestimated values are
473 located upstream of the source location. This discrepancy is mainly due to the lack of information in
474 this portion of the field. Moreover, some concentration values are overestimated particularly in
475 Scenario 2, as a result of the assimilation of only apparent resistivity data and the absence of
476 concentration data. Adding concentration information mitigates this issue, as evidenced by the
477 improved estimation of maximum concentration in Scenario 3.
478 Following a thorough examination of the results and associated metrics, the third scenario, which
479 employs both apparent resistivity data and concentration values, is the best configuration in terms of
480 estimation values and pattern distribution. This comprehensive evaluation emphasizes the importance
481 of integrating different datasets in hydrological studies to achieve a more accurate and reliable
482 estimation of contaminant plume distribution.
483
484 **Table 5** Performance of the proposed approach evaluated for each scenario

Scenario	1	2	3	4	5
ME (mg/L)	-0.08	-0.48	0.06	-0.01	-0.38
MAE (mg/L)	1.99	2.64	1.56	1.97	2.21
RMSE (mg/L)	3.69	3.69	2.74	3.01	3.17
Max Concentration (mg/L)	22.34	32.59	22.00	24.04	28.62
R^2	0.69	0.68	0.82	0.79	0.76
Mean standard deviation (mg/L)	12.51	2.42	1.16	1.43	2.10
Max standard deviation (mg/L)	25.82	13.62	6.18	6.59	13.21

485

486



487

488 **Fig. 5** Estimated concentration distribution for the five scenarios. The injector is marked by a cross.
489 Red circles represent the locations of observation wells

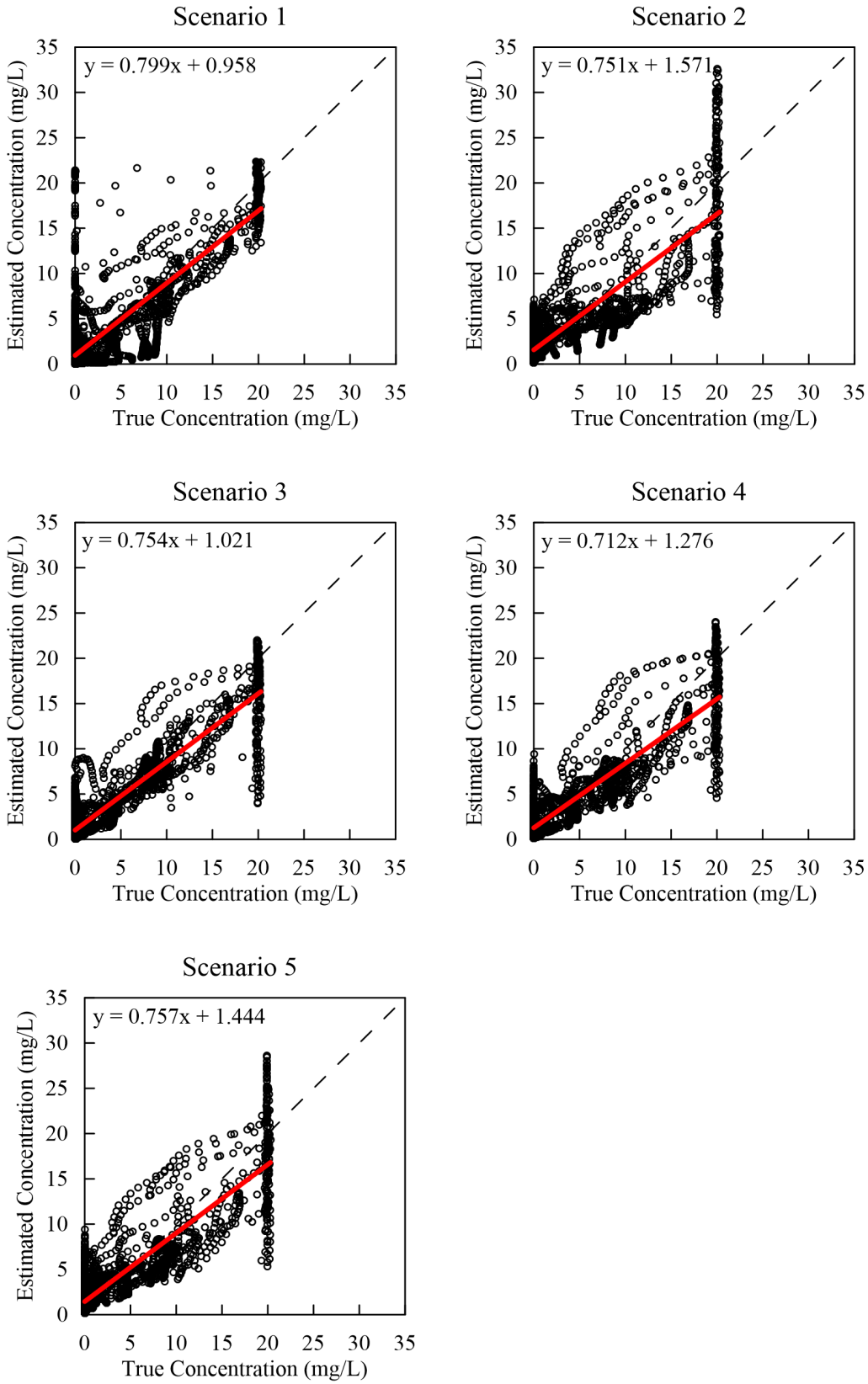


Fig 6 Scatterplot of the estimated vs. observed concentration for all scenarios. The red line is the best linear fit and the black dashed line is the 1:1 line

Fig. 7 represents the agreement between observed values and the corresponding predictions, given by the ensemble mean of the last iteration, for Scenario 2 to Scenario 5. The inclusion of a 45° line serves as a visual benchmark, indicating a perfect fit between the observed and estimated apparent resistivities and concentrations. The proximity of data points to this line signifies the accuracy of the model in reproducing the measured values. The results for Scenario 1 are not explicitly shown as all the points align along the 45° line, since kriging is an exact estimator.

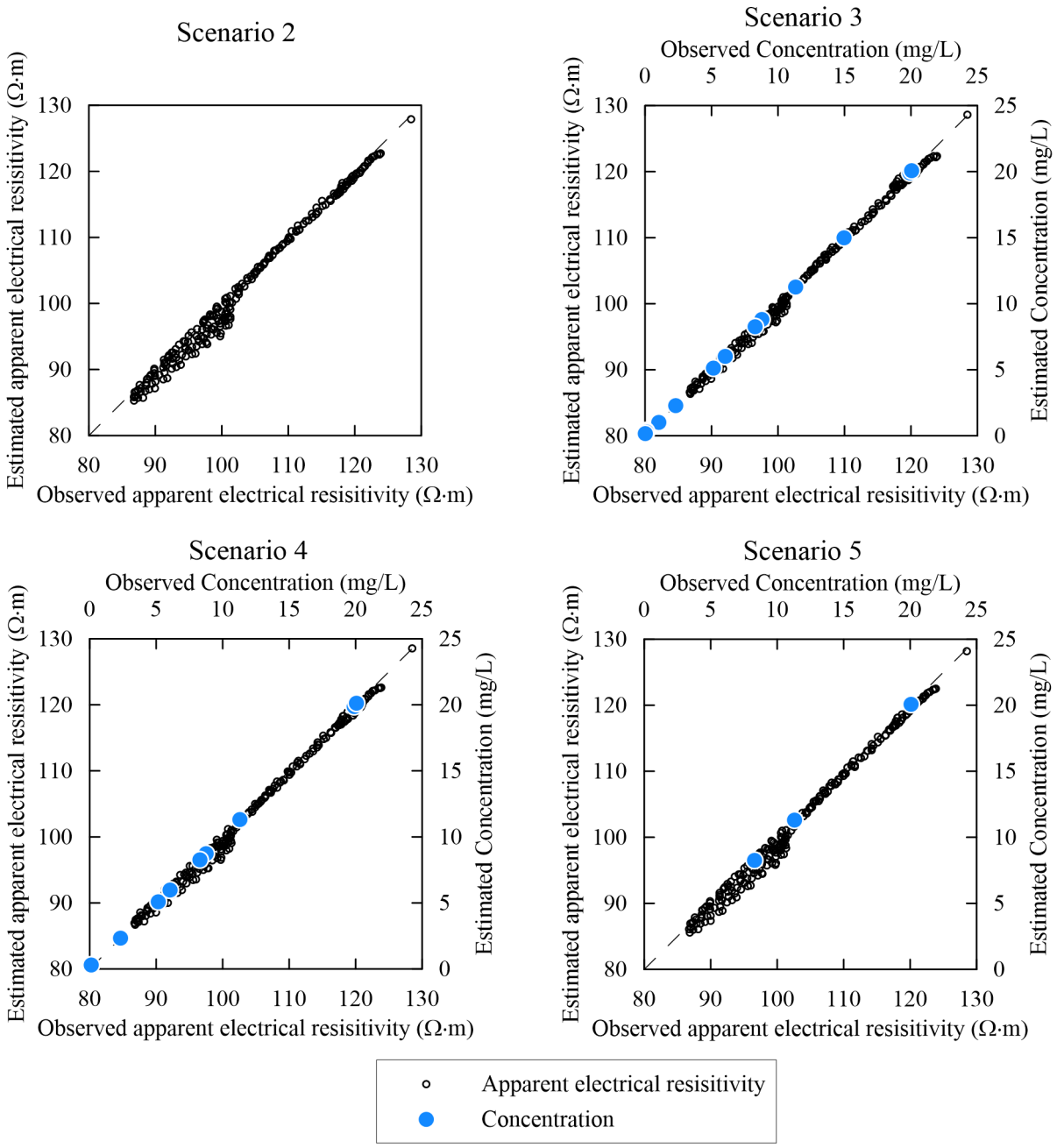


Fig. 7 Observed-Estimated apparent resisitivity and concentration for scenarios 2-5

The uncertainty assessment in the estimation of the concentration map is crucial for a comprehensive understanding of the reliability and robustness of the proposed approach. In this study, the standard deviation serves as a key indicator of the dispersion, or variability, of the estimated concentration maps around their mean (Fig. 8 and Table 5). In Scenario 1, the kriging standard deviation is zero at the observation points and it increases with distance from these points, reaching a maximum of 25.82 mg/L at the borders of the model and the mean of the standard deviation map is 12.51 mg/L. In the remaining scenarios, the standard deviation is computed from the ensemble of the concentration maps. In Scenarios 2 and 5, the standard deviation is high close to the source location where no concentration values are available. Scenario 2 presents a mean value of 2.42 mg/L and a maximum one of 13.62 mg/L. These values are comparable to those in Scenario 5, where the mean and maximum value are 2.10 mg/L and 13.21 mg/L, respectively. Scenario 3 shows the smallest standard deviations with an average value of 1.16 mg/L and a maximum one of 6.18 mg/L. Scenario 4 has a mean (1.43 mg/L) and maximum (6.59 mg/L) values close to those of Scenario 3. The scenario color bar is the same for easy comparison of standard deviation values but is limited to 15 mg/l to optimize the display of Scenarios 3 and 4. In particular, the maximum standard deviation value achieved in Scenario 1 exceeds 25 mg/l, while Scenarios 3 and 4 have values below 7 mg/l. This discrepancy is attributed to Scenario 1 having significantly higher values in the border area, where no concentration information was available.

These results confirm that the combination of ERT and concentration data provides a reliable estimation of the concentration distribution in aquifer. Obviously, the more concentration data there is, the better the result, but even just 3 observations lead to an acceptable result.

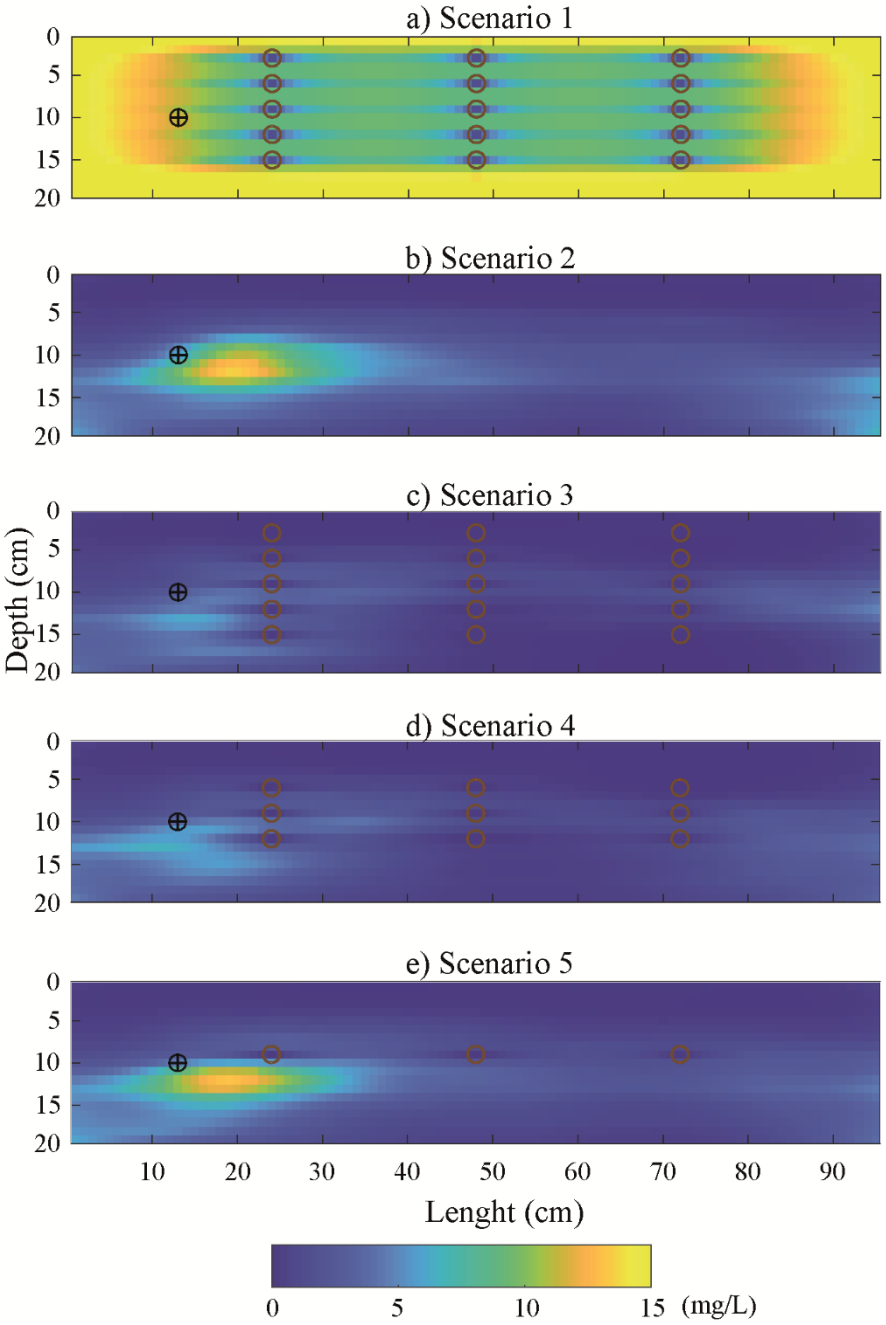
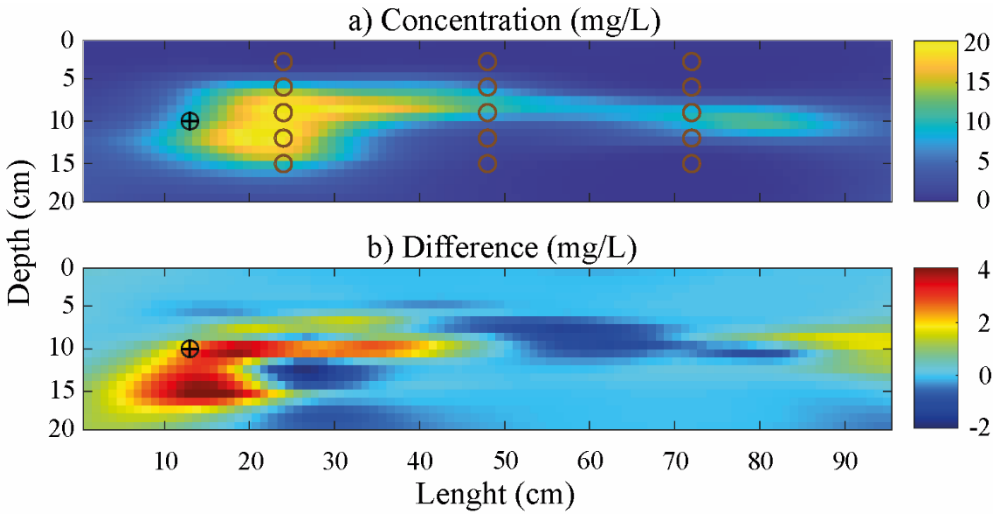


Fig. 4 Standard deviation maps of the estimated concentrations for the five scenarios. The cross denotes the injector. The observation wells are visually depicted by the red circles

4.1 Full forward model (SimPEG) vs CNN

The validity of the proposed inversion approach is further investigated by solving Scenario 3, employing the full forward model instead of the CNN. In Fig. 9a, the estimated plume using the SimPEG forward model is depicted. Fig. 9b illustrates the differences in concentration values

between the two approaches. Remarkably, differences are negligible except for a small area beneath the source location. The two forward models demonstrate comparable performance in solving the inverse problem across several metrics. Both models showcase RMSE values that are close. The full forward model achieves an RMSE of 2.93 mg/L, while the CNN model slightly outperforms it with an RMSE of 2.75 mg/L, suggesting near-equivalent accuracy in predicting the target variable. Furthermore, the full forward model achieves an R^2 of 0.81, closely followed by the CNN model with an R^2 of 0.83. Examining the ME and MAE metrics, which gauge the average magnitude of prediction errors, the full forward model exhibits an ME of 0.32 mg/L and an MAE of 1.68 mg/L, while the CNN model showcases an ME of 0.07 mg/L and an MAE of 1.56 mg/L. Delving into the mean and maximum standard deviation, the CNN model marginally outperforms the full model with slightly lower values for both mean standard deviation (1.11 mg/L vs. 1.68 mg/L) and maximum standard deviation (11.39 mg/L vs. 16.41 mg/L). Finally, both models show similar maximum concentration values, with the CNN model slightly higher at 22.35 mg/L compared to 20.37 mg/L for the full forward model. Notably, a significant disparity arises in terms of computational time: the inverse procedure with the SimPEG forward model takes approximately 2 hours, whereas the one with CNN completes the task in approximately 5 minutes ran with a system composed of an Intel i9-10920X 3.5GHz equipped with 32 GB RAM.



552

553 **Fig 9** a) Estimated concentration distribution, Scenario 3 – SimPEG forward model. The injector is
554 marked by a cross. Red circles represent the locations of observation wells. b) Differences between
555 estimated concentrations (CNN-SimPEG forward model)

556

557 ## 5. Conclusion

558 The presented paper investigated the effectiveness of the Ensemble Smoother with Multiple Data
559 Assimilation (ES-MDA) model in addressing the complex challenge of accurately estimate the
560 spatial distribution of a concentration plume. This is achieved through the simultaneous assimilation
561 of observed electrical resistivity tomography (ERT) data and scattered concentration values from
562 monitoring wells. One of the distinguishing features of this approach was the integration of
563 convolutional neural network (CNN) to speed up the forward model.

564 The study compared five different datasets to evaluate the performance of the proposed approach.
565 These various scenarios enable a thorough examination of the advantages of combining data from
566 multiple sources (Linde and Doetsch 2016), highlighting the effects of different observation
567 datasets on the accuracy of plume distribution assessments. The first scenario used a kriging-based
568 approach to interpolate 15 concentration values, while subsequent scenarios were conducted to
569 evaluate the capability of the proposed inverse hydrogeophysical approach. The second scenario
570 used only apparent resistivity data as observations into the ES-MDA; and the third to fifth scenario
571 combined apparent resistivity data with different subsets of concentration values: 15, 9, and 3,

572 respectively. The third scenario, which combines apparent resistivity with 15 concentration values,
573 emerged as the most promising configuration in terms of accuracy and precision. The least accurate
574 estimates were observed in the case of kriging interpolation (Scenario 1) and ES-MDA utilizing
575 only apparent resistivity data (Scenario 2). A pertinent point to mention, based on the comparison of
576 these results, is the inherent difficulty in relying solely on 15 concentration values derived from a
577 survey for interpolation purposes. This challenge becomes even more pronounced with the use of 9
578 or 3 values, which are insufficient for constructing the variogram in the case of kriging. These
579 findings suggest that such a limited dataset may not provide sufficient information to capture the
580 spatial variability of subsurface concentration maps accurately, emphasizing the importance of
581 combining multiple data sources.
582 In addition, the comparison between the full ERT forward model (i.e., SimPEG) and the CNN
583 showcased significant enhancements in computational efficiency using the surrogate model while
584 maintaining robust predictive performance. The overall results demonstrate the efficacy of the
585 proposed inverse methodology in accurately capturing and predicting the plume concentration's
586 distribution and values, providing a quick tool for supporting optimal strategies for contaminated
587 site remediation.
588 Considering the factors that influence the accuracy of the results, one must keep in mind the
589 petrophysical relationships that play a key role in determining the reliability of concentration
590 estimates. These models may face some uncertainties that might have an impact on the inversion
591 outcomes (Linde et al., 2017). Furthermore, in this work, simplifications have been made in the
592 geophysical properties of the electrical model that have to be considered in real cases. Another
593 factor that could affect the results is the setup of the CNN. For this reason, future researches will
594 focus on a comprehensive analysis of the influence of CNN parameters and hyperparameters on the
595 inversion procedure. Additionally, upcoming works will explore the potential application of the
596 proposed inverse methodology in laboratory experiments.
597

Author Contributions

All authors contributed to the study conception and design. Material preparation, methodology, and analysis were performed by Camilla Fagandini, Valeria Todaro and Andrea Zanini. The first draft of the manuscript was written by Camilla Fagandini and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declaration of Competing Interest

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

Acknowledgments

We thank the Editor and two anonymous Reviewers for their helpful comments. Valeria Todaro acknowledges financial support from PNRR MUR project ECS_00000033_ECOSISTER. Leonardo Azevedo acknowledges the support of CERENA (FCT-UIDB/04028/2020)

References

Alzubaidi L, Zhang J, Humaidi AJ, et al (2021) Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8:. https://doi.org/10.1186/s40537-021-00444-8

Anderson, MP, Woessner, WW (1992) Applied Groundwater Modeling— Simulation of Flow and Advective Transport. Academic Press, Inc., San Diego, CA, 381 p.

Archie, GE (1942) The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics. Transactions of the AIME, 146, 54-62

Asher MJ, Croke BFW, Jakeman AJ, Peeters LJM (2015) A review of surrogate models and their application to groundwater modeling. Water Resour Res 51:5957–5973. https://doi.org/10.1002/2015WR016967

Bear J (1972) Dynamics of fluids in porous media. New York: American Elsevier Pub. Co..

Butera I, Tanda MG, Zanini A (2013) Simultaneous identification of the pollutant release history and the source location in groundwater by means of a geostatistical approach. Stochastic Environmental Research and Risk Assessment 27:1269–1280. https://doi.org/10.1007/s00477-012-0662-1

Camporese M, Cassiani G, Deiana R, et al (2015) Coupled and uncoupled hydrogeophysical inversions using ensemble Kalman filter assimilation of ERT-monitored tracer test data. Water Resour Res 51:3277–3291. https://doi.org/10.1002/2014WR016017

Camporese M, Cassiani G, Deiana R, Salandin P (2011) Assessment of local hydraulic properties from electrical resistivity tomography monitoring of a three-dimensional synthetic tracer test experiment. Water Resour Res 47:. https://doi.org/10.1029/2011WR010528

Carrera J, Alcolea A, Medina A, et al (2005) Inverse problem in hydrogeology. Hydrogeol J 13:206–222

Carpenter P, Ding A, Cheng L. Identifying groundwater contamination using resistivity surveys at a landfill near Maoming. Nat Educ Knowl. 2012;3(7). Available:

https://www.nature.com/scitable/knowledge/library/identifying-groundwater-contamination-using-resistivity-surveys-at-87665012/#:~:text=Resistivity (accessed 22/01/2024)

Chen Y, Zhang D (2006) Data assimilation for transient flow in geologic formations via ensemble Kalman filter. Adv Water Resour 29:1107–1122. https://doi.org/10.1016/j.advwatres.2005.09.007

Chen Z, Gómez-Hernández JJ, Xu T, Zanini A (2018) Joint identification of contaminant source and aquifer geometry in a sandbox experiment with the restart ensemble Kalman filter. J Hydrol (Amst) 564:1074–1084. https://doi.org/10.1016/j.jhydrol.2018.07.073Chen Z, Xu T, Gómez-Hernández JJ, et al (2023a) Reconstructing the release history of a contaminant source with different precision via the ensemble smoother with multiple data assimilation. J Contam Hydrol 252:. https://doi.org/10.1016/j.jconhyd.2022.104115

Chen Z, Xu T, Gómez-Hernández JJ, Zanini A (2021) Contaminant Spill in a Sandbox with Non-Gaussian Conductivities: Simultaneous Identification by the Restart Normal-Score Ensemble Kalman Filter. Math Geosci 53:1587–1615. https://doi.org/10.1007/s11004-021-09928-y

Chen Z, Zong L, Gómez-Hernández JJ, et al (2023b) Contaminant source and aquifer characterization: An application of ES-MDA demonstrating the assimilation of geophysical data. Adv Water Resour 181:. https://doi.org/10.1016/j.advwatres.2023.104555

Citarella D, Cupola F, Tanda MG, Zanini A (2015) Evaluation of dispersivity coefficients by means of a laboratory image analysis. J Contam Hydrol 172:10–23. https://doi.org/10.1016/j.jconhyd.2014.11.001

Cockett R, Kang S, Heagy LJ, et al (2015) SimPEG: An open source framework for simulation and gradient based parameter estimation in geophysical applications. Comput Geosci 85:142–154

Crestani E, Camporese M, Salandin P (2015) Assessment of hydraulic conductivity distributions through assimilation of travel time data from ERT-monitored tracer tests. Adv Water Resour 84:23–36. https://doi.org/10.1016/j.advwatres.2015.07.022

661 Crestani E, Camporese M, Bau D, et al (2013) Ensemble Kalman Filter Versus Ensemble Smoother
662 for Assessing Hydraulic Conductivity via Tracer Test Data Assimilation. Hydrol Earth Syst Sci
663 17:1517–1531. https://doi.org/10.5194/hessd-9-13083-2012
664 Cupola F, Tanda MG, Zanini A (2015) Laboratory sandbox validation of pollutant source location
665 methods. Stochastic Environmental Research and Risk Assessment 29:169–182.
666 https://doi.org/10.1007/s00477-014-0869-4
667 Das V, Pollack A, Wollner U, Mukerji T (2019) Convolutional neural network for seismic impedance
668 inversion. Geophysics 84:R869–R880. https://doi.org/10.1190/geo2018-0838.1
669 D'Oria M, Mignosa P, Tanda MG, Todaro V (2022) Estimation of levee breach discharge
670 hydrographs: comparison of inverse approaches. Hydrological Sciences Journal 67:54–64.
671 https://doi.org/10.1080/02626667.2021.1996580
672 Edwards LS (1977) A modified pseudosection for resistivity and induced-polarization. Geophysics,
673 42, 1020-1036.
674 El Idrysy EH, De Smedt F (2007) A comparative study of hydraulic conductivity estimations using
675 geostatistics. Hydrogeol J 15:459–470. https://doi.org/10.1007/s10040-007-0166-0
676 Emerick AA, Reynolds AC (2012) History matching time-lapse seismic data using the ensemble
677 Kalman filter with multiple data assimilations. Comput Geosci 16:639–659
678 Emerick AA, Reynolds AC (2013) Ensemble smoother with multiple data assimilation. Comput
679 Geosci 55:3–15. https://doi.org/10.1016/j.cageo.2012.03.011
680 Evensen G (1994) Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte
681 Carlo methods to forecast error statistics. J Geophys Res 99:. https://doi.org/10.1029/94jc00572
682 Fernandez Visentini A, Linde N, Borgne T Le, Dentz M (2020) Inferring geostatistical properties of
683 hydraulic conductivity fields from saline tracer tests and equivalent electrical conductivity time-
684 series. Adv Water Resour 146:. https://doi.org/10.1016/j.advwatres.2020.103758
685 Frohlich RK, Urish DW (2002) The use of geoelectrics and test wells for the assessment of
686 groundwater quality of a coastal industrial site

687 Godoy VA, Napa-García GF, Gómez-Hernández JJ (2022) Ensemble Smoother with Multiple Data
688 Assimilation as a Tool for Curve Fitting and Parameter Uncertainty Characterization: Example
689 Applications to Fit Nonlinear Sorption Isotherms. Math Geosci 54:807–825.
690 https://doi.org/10.1007/s11004-021-09981-7
691 Gómez-Hernández JJ, Hendricks Franssen HJ, Sahuquillo A (2003) Stochastic conditional inverse
692 modeling of subsurface mass transport: a brief review and the self-calibrating method. Stochastic
693 Environmental Research and Risk Assessment, 17(5), 319-328
694 Gómez-Hernández JJ, Wen XH (1994) Probabilistic assessment of travel times in groundwater
695 modeling. Stochastic Hydrology and Hydraulics, 8, 19-55.
696 Gómez-Hernández JJ, Xu T (2022) Contaminant Source Identification in Aquifers: A Critical View.
697 Math Geosci 54:437–458. https://doi.org/10.1007/s11004-021-09976-4
698 Harbaugh, AW (2005) MODFLOW-2005, the US Geological Survey Modular Ground-Water Model:
699 The Ground-Water Flow Process, Vol. 6. US Department of the Interior, US Geological Survey,
700 Reston, VA, USA.
701 Gu J, Wang Z, Kuen J, et al (2018) Recent advances in convolutional neural networks. Pattern
702 Recognit 77:354–377. https://doi.org/10.1016/j.patcog.2017.10.013
703 Hendricks Franssen HJ, Alcolea A, Riva M, Bakr M, Van der Wiel N, Stauffer F, Guadagnini A
704 (2009) A comparison of seven methods for the inverse modelling of groundwater flow. Application
705 to the characterisation of well catchments. Adv Water Resour 32:851–872.
706 https://doi.org/10.1016/j.advwatres.2009.02.011
707 Hong J, Liu J. (2020). Rapid estimation of permeability from digital rock using 3D convolutional
708 neural network. Comput Geosci. 24: 1523–1539. doi:10.1007/s10596-020-09941-w
709 Hubbard S. (2011) Hydrogeophysics. Lawrence Berkeley National Laboratory. Available on
710 https://escholarship.org/uc/item/11c8s8d4 (accessed 22/01/2024)

711 Huysmans M, Dassargues A (2009) Application of multiple-point geostatistics on modelling
712 groundwater flow and transport in a cross-bedded aquifer (Belgium). Hydrogeol J 17:1901–1911.
713 https://doi.org/10.1007/s10040-009-0495-2
714 Irving J, Singha K (2010) Stochastic inversion of tracer test and electrical geophysical data to estimate
715 hydraulic conductivities. Water Resour Res 46:. https://doi.org/10.1029/2009WR008340
716 Jamshidi A, Samani JMV, Samani HMV, et al (2020) Solving inverse problems of unknown
717 contaminant source in groundwater-river integrated systems using a surrogate transport model based
718 optimization. Water (Switzerland) 12:. https://doi.org/10.3390/w12092415
719 Kang X, Shi X, Deng Y, et al (2018) Coupled hydrogeophysical inversion of DNAPL source zone
720 architecture and permeability field in a 3D heterogeneous sandbox by assimilation time-lapse cross-
721 borehole electrical resistivity data via ensemble Kalman filtering. J Hydrol (Amst) 567:149–164.
722 https://doi.org/10.1016/j.jhydrol.2018.10.019
723 Kang X, Shi X, Revil A, et al (2019) Coupled hydrogeophysical inversion to identify non-Gaussian
724 hydraulic conductivity field by jointly assimilating geochemical and time-lapse geophysical data. J
725 Hydrol (Amst) 578:. https://doi.org/10.1016/j.jhydrol.2019.124092
726 Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep
727 convolutional neural networks. Artif Intell Rev 53:5455–5516. https://doi.org/10.1007/s10462-020-
728 09825-6
729 Kitanidis PK (1995) Quasi-Linear Geostatistical Theory for Inversing. Water Resour Res 31:2411–
730 2419. https://doi.org/10.1029/95WR01945
731 Lähivaara T, Malehmir A, Pasanen A, et al (2019) Estimation of groundwater storage from seismic
732 data using deep learning. Geophys Prospect 67:2115–2126. https://doi.org/10.1111/1365-2478.12831
733 LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document
734 recognition. Proceedings of the IEEE, 86(11), 2278-2324
735 Li L, Zhou H, Hendricks Franssen HJ, Gómez-Hernández JJ (2012) Modeling transient groundwater
736 flow by coupling ensemble Kalman filtering and upscaling. Water Resources Research, 48(1).

737 Li J, Lu W, Wang H, Fan Y (2019) Identification of groundwater contamination sources using a
738 statistical algorithm based on an improved Kalman filter and simulation optimization. Hydrogeol J
739 27:2919–2931. https://doi.org/10.1007/s10040-019-02030-y
740 Linde N, Doetsch J (2016) Joint inversion in hydrogeophysics and near-surface geophysics. In
741 Integrated imaging of the earth (pp. 117–135). John Wiley & Sons, Inc.
742 https://doi.org/10.1002/9781118929063.ch7
743 Linde N, Ginsbourger D, Irving J, et al (2017) On uncertainty quantification in hydrogeology and
744 hydrogeophysics. Adv Water Resour 110:166–181. https://doi.org/10.1016/j.advwatres.2017.10.014
745 Loke MH, Chambers JE, Rucker DF, et al (2013) Recent developments in the direct-current
746 geoelectrical imaging method. J Appl Geophy 95:135–156.
747 https://doi.org/10.1016/j.jappgeo.2013.02.017
748 Mavko G, Mukerji T, Dvorkin J (2009) The Rock Physics Handbook. Cambridge University Press.
749 doi:10.1017/CBO9780511626753
750 Michalak AM, Kitanidis PK (2004) Estimation of historical groundwater contaminant distribution
751 using the adjoint state method applied to geostatistical inverse modeling. Water Resour Res 40:.
752 https://doi.org/10.1029/2004WR003214
753 McLaughlin D, Townley LR (1996) A reassessment of the groundwater inverse problem. Water
754 Resour Res 32:1131–1161
755 Page LM (1969) The use of the geoelectric method for investigating geologic and hydrologic
756 conditions in Santa Clara County, California. J Hydrol. 7: 167–177. doi:10.1016/0022-
757 1694(69)90054-7
758 Panahi M, Sadhasivam N, Pourghasemi HR, et al (2020) Spatial prediction of groundwater potential
759 mapping based on convolutional neural network (CNN) and support vector regression (SVR). J
760 Hydrol (Amst) 588:. https://doi.org/10.1016/j.jhydrol.2020.125033
761 Pereira JL, Gómez-Hernández JJ, Zanini A, et al (2023) Iterative geostatistical electrical resistivity
762 tomography inversion. Hydrogeol J 31:1627–1645. https://doi.org/10.1007/s10040-023-02683-w

763 Pidlisecky A, Knight R (2008) FW2_5D: A MATLAB 2.5-D electrical resistivity modeling code.
764 Comput Geosci 34:1645–1654. https://doi.org/10.1016/j.cageo.2008.04.001
765 Pollock D, Cirpka OA (2010) Fully coupled hydrogeophysical inversion of synthetic salt tracer
766 experiments. Water Resour Res 46:. https://doi.org/10.1029/2009WR008575
767 Pollock D, Cirpka OA (2012) Fully coupled hydrogeophysical inversion of a laboratory salt tracer
768 experiment monitored by electrical resistivity tomography. Water Resour Res 48:.
769 https://doi.org/10.1029/2011WR010779
770 Puzyrev V (2012) Deep learning electromagnetic inversion with convolutional neural networks.
771 https://doi.org/10.1093/gji/ggz204/5484841
772 Pyrcz M, Jo H, Kupenko A, Liu W, Gigliotti A E, Salomaki T, Javier S (2021) GeostatsPy:
773 Geostatistical Library in Python (Version 1.0.0) [Computer software].
774 https://doi.org/10.5281/zenodo.
775 Rubin Y, Hubbard SS (2005) Hydrogeophysics. Dordrecht: Springer Netherlands; doi:10.1007/1-
776 4020-3102-5
777 Secci D, Molino L, Zanini A (2022) Contaminant source identification in groundwater by means of
778 artificial neural network. J Hydrol (Amst) 611:. https://doi.org/10.1016/j.jhydrol.2022.128003
779 Secci D, A. Godoy V, Gómez-Hernández JJ (2024) Physics-Informed Neural Networks for solving
780 transient unconfined groundwater flow. Comput Geosci 182:.
781 https://doi.org/10.1016/j.cageo.2023.105494
782 Tang M, Liu Y, Durlofsky LJ (2021) Deep-learning-based surrogate flow modeling and geological
783 parameterization for data assimilation in 3D subsurface flow. Comput Methods Appl Mech Eng
784 376:113636. doi: 10.1016/j.cma.2020.113636
785 Todaro V, D'Oria M, Tanda MG, Gómez-Hernández JJ (2019) Ensemble smoother with multiple
786 data assimilation for reverse flow routing. Comput Geosci 131:32–40.
787 https://doi.org/10.1016/j.cageo.2019.06.002

788 Todaro V, D'Oria M, Tanda MG, Gómez-Hernández JJ (2021) Ensemble smoother with multiple
789 data assimilation to simultaneously estimate the source location and the release history of a
790 contaminant spill in an aquifer. J Hydrol (Amst) 598:. https://doi.org/10.1016/j.jhydrol.2021.126215
791 Todaro V, D'Oria M, Tanda MG, Gómez-Hernández JJ (2022) genES-MDA: A generic open-source
792 software package to solve inverse problems via the Ensemble Smoother with Multiple Data
793 Assimilation. Comput Geosci 167:. https://doi.org/10.1016/j.cageo.2022.105210
794 Todaro V, D'Oria M, Zanini A, et al (2023) Experimental sandbox tracer tests to characterize a two-
795 facies aquifer via an ensemble smoother. Hydrogeol J 31:1665–1678. https://doi.org/10.1007/s10040-
796 023-02662-1
797 Van Leeuwen PJ, Evensen G (1996) Data assimilation and inverse methods in terms of a probabilistic
798 formulation. Mon Weather Rev 124:2898–2913. https://doi.org/10.1175/1520-
799 0493(1996)124<2898:DAAIMI>2.0.CO;2
800 Vereecken H, Binley A, Cassiani G, Revil A, Titov K (2006) Applied Hydrogeophysics. Dordrecht:
801 Springer Netherlands; pp. 1–8. doi:10.1007/978-1-4020-4912-5_1
802 Vishal Das, Ahinoam Pollack, Uri Wollner, and Tapan Mukerji, (2019), "Convolutional neural
803 network for seismic impedance inversion," GEOPHYSICS 84: R869-R880
804 Visentini AF, Linde N, Borgne T Le, Dentz M (2020) Inferring geostatistical properties of hydraulic
805 conductivity fields from saline tracer tests and equivalent electrical conductivity time-series. Adv
806 Water Resour 146:. https://doi.org/10.1016/j.advwatres.2020.103758
807 Vladimir Puzyrev, Deep learning electromagnetic inversion with convolutional neural networks,
808 Geophysical Journal International, Volume 218, Issue 2, August 2019, Pages 817–832,
809 https://doi.org/10.1093/gji/ggz204
810 Wilson SR, Ingham M, McConchie JA (2006) The applicability of earth resistivity methods for saline
811 interface definition. J Hydrol (Amst) 316:301–312. https://doi.org/10.1016/j.jhydrol.2005.05.004

812 Xu T, Gómez-Hernández JJ (2016) Joint identification of contaminant source location, initial release
813 time, and initial solute concentration in an aquifer via ensemble Kalman filtering. Water Resour Res
814 52:6587–6595. https://doi.org/10.1002/2016WR019111
815 Xu T, Gómez-Hernández JJ (2018) Simultaneous identification of a contaminant source and hydraulic
816 conductivity via the restart normal-score ensemble Kalman filter. Adv Water Resour 112:106–123.
817 https://doi.org/10.1016/j.advwatres.2017.12.011
818 Xu T, Gómez-Hernández JJ, Chen Z, Lu C (2021) A comparison between ES-MDA and restart EnKF
819 for the purpose of the simultaneous identification of a contaminant source and hydraulic conductivity.
820 J Hydrol (Amst) 595. https://doi.org/10.1016/j.jhydrol.2020.125681
821 Zanini A, Woodbury AD (2016) Contaminant source reconstruction by empirical Bayes and Akaike's
822 Bayesian Information Criterion. J Contam Hydrol 185–186:74–86.
823 https://doi.org/10.1016/j.jconhyd.2016.01.006
824 Zhang Z, (2018) Improved Adam Optimizer for Deep Neural Networks. IEEE/ACM 26th
825 International Symposium on Quality of Service (IWQoS), Banff, AB, Canada, 2018, pp. 1-2, doi:
826 10.1109/IWQoS.2018.8624183.
827 Zhou H, Gómez-Hernández JJ, Li L (2012) A pattern-search-based inverse method. Water Resources
828 Research, 48:3.
829 Zhou H, Gómez-Hernández JJ, Li L (2014) Inverse methods in hydrogeology: Evolution and recent
830 trends. Adv Water Resour 63:22–37
831 Zheng C, Wang PP. MT3DMS: A modular three-dimensional multispecies transport model for
832 simulation of advection, dispersion, and chemical reactions of contaminants in groundwater
833 systems. Vicksburg, MS, USA: U.S. Army Engineer Research and Development Center No.
834 SERDP-99-1; 1999.
835 Zhou Z, Zabaras N, Tartakovsky DM, (2022) Deep Learning for Simultaneous Inference of
836 Hydraulic and Transport Properties. Water Resour. Res. 58. doi: 10.1029/2021WR031438

837 Zimmerman DA, De Marsily G, Gotway CA, et al (1998) A comparison of seven geostatistically
838 based inverse approaches to estimate transmissivities for modeling advective transport by
839 groundwater flow. Water Resour Res 34:1373–1413. https://doi.org/10.1029/98WR00003