Mapping of hydraulic transmissivity field from inversion of tracer test data using convolutional neural networks. CNN-2T

M. T. Vu, Abderrahim Jardani

▶ To cite this version:

M. T. Vu, Abderrahim Jardani. Mapping of hydraulic transmissivity field from inversion of tracer test data using convolutional neural networks. CNN-2T. Journal of Hydrology, 2022, 606, ⟨10.1016/j.jhydrol.2022.127443⟩. ⟨insu-03661810⟩

HAL Id: insu-03661810
https://insu.hal.science/insu-03661810v1
Submitted on 22 Jul 2024

Mapping of hydraulic transmissivity field from inversion of tracer test data using convolutional neural networks. CNN-2T

M. T. Vu [1*], A. Jardani [1]

[1] Université de Rouen, M2C, UMR 6143, CNRS, Morphodynamique Continentale et Côtière, GeoDeeepLearning Consortium, Mont Saint Aignan, France

* Corresponding author: minh-tan.vu@univ-rouen.fr

Emails: abderrahim.jardani@univ-rouen.fr; minh-tan.vu@univ-rouen.fr;

ABSTRACT

This paper introduces a new concept for mapping hydraulic transmissivity from temporal concentration data collected in multiple tracer tests. Based on convolutional neural network, the principle uses an encoder-decoder architecture with multiple neural layers to establish a relationship between the concentration data and the transmissivity field. This relationship is established in two phases with two networks. The first network is designated and trained to reconstruct a transmissivity field using data from a single tracer test. To improve the reconstruction quality, the second network then performs a joint interpretation for multiple tracer tests, which reprocesses all the transmissivity resulted from the first network for each individual tracer test. Both networks are trained by synthetic data, where the transmissivity models are generated with a Gaussian variogram and its properties are considered as prior information on the aquifer heterogeneity. Tracer tests are derived numerically by solving the forward problem to obtain the corresponding concentration data that feed the training. The trained networks accurately map the transmissivity fields, of which the accuracy relies on the volume and nature of the heterogeneities of training models, as well as the number of piezometers used to monitor the concentration changes. Reconstruction quality, on the other hand, is less influenced by data noise. Effective training requires a large dataset, but the time required for dataset generation is only on the order of the Gauss-Newton algorithm in a conventional inversion, while the trained network performs inference instantly.

Keywords: Inversion method, Hydraulic tomography, CNN, Neural network architecture, Deep Learning, Tracer test.

1. Introduction

Characterising the spatial variability of aquifer hydraulic properties in hydrogeology remains a complex and critical task for any attempt to understand water dynamics and contaminant transport in aquifers (Bear & Cheng, 2010). Spatial heterogeneity, particularly in hydraulic conductivity, drives transport mechanism of water and contaminants in the subsurface. Such characterization, referred as hydraulic tomography, is typically performed on a large scale by analysing hydraulic changes due to natural forces, such as the hydrological cycle and tidal oscillations (Bailey & Baù, 2012; Jardani et al. 2012). On a small scale, the task can be similarly achieved through a joint interpretation of hydraulic head variations in a sequential multiple pumping test (Yeh & Liu, 2000; Cardiff et al., 2013). As an alternative, tracer tests can provide insight into hydraulic properties, particularly in high permeability environments, by analysing the temporal concentration variations of an injected tracer, i.e., a dye, salt, or heat (Yeh & Zhu, 2007; Jardani et al., 2013; Lee & Kitanidis, 2014; Sanchez-León, 2016).

To interpret these hydraulic/concentration measurements, an inversion solver often involves an iterative algorithm that fits the measurement from adjusting a hydraulic conductivity distribution. However, this type of inversion approach is ill-posed, resulting in a non-unique solution. To overcome this issue, constraints and prior information must be included which narrow the possibilities and promote convergence to a physically meaningful model (Kitanidis, 1997; Zha et al., 2017).

This optimization can be performed with an algorithm that belongs to one of three types: deterministic, stochastic, and global (Sen et al. 1995; Carrera et al. 2005; Soueid et al. 2015; Xu & Gómez-Hernández, 2018; Reuschen et al., 2020). In deterministic algorithms, such as Gauss-Newton and Conjugate Gradient, optimization aims at a local minimum by iteratively computing the gradient of the objective function. Their computation time and accuracy rely on the selection of initial model and the calibration of Jacobian matrix (Tarantola & Valette, 1982; Fahs et al., 2014). Stochastic algorithms, otherwise, search for the best model, which is based on a random strategy in which many models are generated, selected, or discarded according to their ability to fit both the water/concentration data and the prior information (Cui & O'sullivan, 2011). This type of algorithm is straightforward to implement, and the resulting model is independent of the initial choice. However, the computation demands an intensive resource as it requires solving several times the forward problem. Global algorithms, such as Simulated Annealing and Genetic, also represent an interesting alternative to invert the hydraulic/concentration data by iteratively searching for the global optimum (Datta-Gupta et al, 1995; Jha & Datta, 2011).

Among the applications of these inversion tools on the reconstruction of hydraulic parameters by processing tracer test data, which will be the main subject of this paper, Zhu et al. (2008) showed a synthetic case that interpretation of tracer test data with a linear sequence estimator. The approach provided useful information about the heterogeneity of the hydraulic conductivity field in a short time, especially when the concentration data are expressed with their temporal moments. Schewde et al. (2013) employed the quasi-linear geostatistical approach to characterize 3D hydraulic conductivity by jointly inverting hydraulic and thermal data derived from pumping and heat tracing tests conducted on a synthetic aquifer. Lee and Kitanidis (2014) applied principal component geostatistical inversion, derived from the classical quasi-linear geostatistical algorithm, to perform a joint inversion of hydraulic test and tracer data. Sanchez-Leon et al. (2016) examined the effectiveness of the Ensemble Kalman Filter and the Kalman Ensemble Generator in interpreting synthetic pumping and tracer test data. Jimènez et al. (2016) opted for the stochastic Reversible-Jump Markov chain Monte Carlo with the pilot point as a parameterization method to invert tracer data collected on alluvial environment. Saley et al. (2016) used the Hamiltonian Monte Carlo algorithm to reconstruct hydraulic conductivity heterogeneity by inverting heat tracing data collected on sandbox. Somogyvári & Bayer (2017) mapped the hydraulic conductivity of a heterogeneous alluvial aquifer by inverting thermal tracer data with the Simultaneous Iteration Reconstruction Technique (SIRT). The same algorithm was tested by Kittilä et al. 2020, in the interpretation of dye tracing test data performed on fractured crystalline rocks to identify preferential groundwater flow paths.

Thus, the literature is rich with applications of inversion tools for tracer data processing; however, an algorithm based on a new generation of deep learning tools has barely been performed for tracer tests. The deep learning algorithm is based on the approximation of an inversion function that non-linearly relates two parameters: the spatio-temporal measurement data and the aquifer properties (Zarita & Ong, 2008). These relationships are determined using networks consists of multiple layers in which a set of linear and nonlinear operations are performed sequentially on the neurons of these layers. This process involves the use of multidimensional coefficients, known as weights and biases, to ensure the connectivity between two successive layers (Zhang et al., 1998; Tahmasebi et al., 2020; Marçais & de Dreuzy, 2017). The identification of the weights and biases is based on the optimization process, i.e. their selection is made according to their ability to connect the input data with their corresponding outputs which are considered to be known and referred to as training dataset (Hassoun, 1995). The accuracy of the predictions provided by the learned network depends strongly on the size of this training dataset (Zhang et al.1998; Jardani et al., 2022).

In the processing of pumping and tracer test data, the first attempts using deep learning algorithms to approximate inversion functions were made with vanilla neural networks, where the layers are fully connected (Zio, 1997; Balkhair, 2002; Lin & Chen, 2006; Akin, 2005; Yoon et al., 2007). This kind of connectivity implies the use of many parameters, which overburdens the learning process and requires a heavy resources and computation time. However, in recent years, a new generation of networks based on convolutional neural networks has emerged, allowing to better address the problems of tomography in geosciences (LeCun & Bengio, 1995; LeCun et al., 1998). Indeed, this type of network permits efficiently linking two images through operating several convolution layers. Convolution is a linear operation in which small weights called filters are applied to the input image to establish local connectivity between neighbouring pixels (Indolia et al., 2018). Each filter is dedicated to extracting a feature in the image by scanning the filter region-wise over the entire image. The use of these small filters allows the recovery of the main features on the input image and reduces the number of unknown parameters learned during training (O'Shea & Nash, 2015).

The number of applications of these new networks in the geosciences continues to grow, including Liu et al. (2019) who used the Unet network to train an inversion operator for electrical resistivity tomography that links apparent electrical resistivity data collected at the ground surface to the distribution of electrical resistivity in the subsurface. Vu & Jardani (2021) trained a convolutional neural network with encoder-decoder architecture to perform 3D inversion of electrical resistivity data. Apolinario et al. (2019) opted for the convolutional neural networks with encoder-decoder structure to estimate seismic velocity from acoustic data. Zhang & Lin (2020) employed the

Generative Adversarial Network to approximate the inversion function to image spatial heterogeneity of seismic velocity for seismogram data. In the electromagnetic applications, Puzyrev & Swidinsky (2019) used a convolution neural network to perform 1D inversion of electromagnetic data to obtain the vertical profile of electrical conductivity. In the hydraulic tomography applications, Bao et al. (2020) combined Ensemble Smoother with Multiple Data Assimilation with Generative Adversarial Networks to map the geometry of the channels by analysing the hydraulic head and concentration data. Sun (2018) used the Generative Adversarial Networks to identify the spatial heterogeneity of the synthetic aquifer for the pumping test data. Laloy et al (2018) also used an inversion of hydraulic data based on the Generative Adversarial Networks to determine spatial distribution of hydraulic conductivity fields in 2D and 3D. Most of these few initiatives were dedicated to reconstructing the hydraulic conductivity field from hydraulic head observations. However, there has been little work on the inversion of solute data.

In this paper, we investigate the effectiveness of a network of encoder-decoder structures to reconstruct the transmissivity field from temporal concentration data of a salt tracer. This network called SegNet was originally developed for a segmentation task in which networks are trained to identify pixels of labelled objects (animal, lane, bicycle, car) on images (Badrinarayanan et al., 2017). However, in this paper, the network is adapted to perform a regression task to associate concentration data with hydraulic transmissivity fields. To our knowledge, this is the first time such an advanced network has been performed in a tracer test inversion problem.

The first part of this paper introduces the theoretical background of the tracer test - the forward problem - used to generate the training data. The principle of inversion is then introduced, followed by a brief explanation of the proposed encoder-decoder architecture, its built-in layers, and operations for establishing the relationship between the concentration data and hydraulic transmissivity fields. In the implementation process, the network is tested with 5000 models after being trained with synthetic datasets. After that, the algorithm accuracy is assessed under a variety of data conditions in a monitoring field, including the effects of data size, low resolution, and observation noise. The effect of prior conditions introduced into the generation of training models on accuracy is discussed further in a range of real field contexts. Finally, the Deep Learning results are compared with those determined with a Gauss-Newton method to illustrate effectiveness of the proposed algorithm.

2. Theoretical Background

In this section, we first describe the theoretical concept of the forward problem used for the numerical simulation of spatiotemporal concentration data of a tracer injected into an aquifer with a heterogeneous transmissivity field. This forward problem is based on a coupling of the steady-state groundwater flow equation and the transient transport equation and is employed to build a database in which the concentration data and its corresponding transmissivity field are used in the network training phase. In the second section, we introduce the concept of inversion problem with the convolutional neural network with an encoder-decoder structure.

2.1 Forward problem

Numerical simulation of the transfer of a soluble tracer in a saturated heterogeneous medium relies on the sequential solving of the groundwater flow and transport equations. First, the groundwater flow equation is solved to determine the spatial distribution of the Darcy velocity, which is subsequently used in the transport equation to compute the convective and dispersive terms. In this paper, the groundwater flow equation is expressed in the steady state (Bear, 1972):

$$\nabla \cdot (T\nabla h) = q_0 \delta(x - x_0), \tag{1}$$

subjected to a boundary condition

$$h = h_0 \text{ at } \Gamma, \tag{2}$$

where T is the heterogeneous transmissivity field (m^2/s) generated geostatistically, q_0 represents the injection source (m^2/s) at the well position x_0, δ is Kronecker delta. h is the hydraulic head (m). h_0 is a constant hydraulic head imposed at the boundary Γ.

The water velocity is determined by Darcy's law:

$$\boldsymbol{u} = -K\,\nabla h, \tag{3}$$

where, K (m/s) denotes the hydraulic conductivity, $K = T/b$ with b (m) is thickness of the aquifer taken as constant in this study.

Then the velocity field is used in the transport equation that is expressed in transient mode as following:

$$\phi\frac{\partial c}{\partial t} + \nabla\cdot(-\phi\boldsymbol{D}\cdot\nabla c) + \boldsymbol{u}\nabla c = 0, \tag{4}$$

subjected to the following initial and boundary conditions

$$c = 0 \ \text{ at } \ t = 0, \tag{5}$$

$$-\boldsymbol{n}\cdot(\boldsymbol{D}\nabla c) = c_0 \ \text{ at } \ x = x_0, \tag{6}$$

where ϕ (.) is the dynamic porosity, c (mol/m^3) is the solute concentration, t (s) is the time, $\boldsymbol{D}$ (m^2/s) represents the hydrodynamic dispersion tensor which can be given by:

$$\boldsymbol{D} = \left(\frac{D_m}{\tau} + \alpha_T v\right)\boldsymbol{I}_3 + \frac{\alpha_L - \alpha_T}{v}\boldsymbol{u}\otimes\boldsymbol{u}, \tag{7}$$

where D_m (m^2/s) is molecular diffusion coefficient of tracer, v (m/s) is the magnitude of velocity, $v = ||\boldsymbol{u}||$, τ (.) is the tortuosity, α_L and α_T are the longitudinal and transverse dispersivities, $\boldsymbol{I}_3$ is the identify matrix 3 by 3.

In this paper, the groundwater flow and transport equations are solved numerically using the finite element method with Comsol software. These equations constitute the direct problem of the tracer test in which the concentration of conservative tracer is connected in a nonlinear manner to the transmissivity field:

$$\boldsymbol{c} = f(\boldsymbol{m}), \text{with } \boldsymbol{T} = 10^{-m}, \tag{8}$$

As mentioned earlier, the forward problem will be served to create the training dataset by computing concentration data for thousands of geostatistically generated transmissivity fields.

2.2 *Inverse problem with deep learning network*

In this section, we describe the theoretical background of inversion with an encoder-decoder structure neural network that is based on the approximation of the highly nonlinear inverse function relating the concentration data (input) to the transmissivity model (output). This approximation is derived from a statistical analysis of the provided concentration-transmissivity data in which a set of weight and bias parameters are determined by an optimization algorithm to construct a universal relationship between the input and output data. Thus, the inverse operation can be reformulated as following:

$$\boldsymbol{m} = f^{-1}(\boldsymbol{c}, \boldsymbol{\Theta}), \tag{9}$$

where f^{-1} denotes the inverse operator and $\boldsymbol{\Theta}$ denotes all parameters used in the network and are determined in the training step. $\boldsymbol{c}$ is the temporal measurements of the concentration of tracer collected in the wells. $\boldsymbol{m}$ is the negative of logarithm of transmissivity field.

In the learning phase, the inverse operator is built with an optimization process for finding the best weights and biases that connect the concentration and transmissivity of the training data set. This operation can be expressed as follows:

$$\boldsymbol{\Theta} = \mathrm{argmin}\left\{\sum_{1}^{N} \|\boldsymbol{m}_i - f^{-1}(\boldsymbol{c}_i, \boldsymbol{\Theta})\|\right\}, \tag{10}$$

where ($\boldsymbol{mi}$, $\boldsymbol{c}_i$) refers to the transmissivity model and its corresponding concentration data determined numerically with the forward problem, and N is the size of training data used to train the network.

The quality of the approximation of the inversion function depends on the amount of data used in the training phase. To get an accurate estimate, we need to build a large training database from repeated solving of the forward problem, making this step the most time-consuming in the process. The structure of the networks also has a significant impact on the quality of the reconstruction of inversion function. In this paper, we chose the Segnet network, which was initially established to deal with the problem of segmentation on images by identifying the objects learned on the images (such as: car, bus, road, animal). Since it is difficult to perform a soil classification from the concentration data, we adapted this network to handle a regression problem for linking the concentration data to the transmissivity model.

The Segnet network is composed of Encoder and Decoder blocks presented in a symmetrical manner (Badrinarayanan et al., 2017). On each encoder, several operations are performed sequentially to extract the main features of the data using convolution, batch normalization, and ReLu.

Convolution is a linear operation in which several small filters are slid to cover the entire image to recover the relevant information called feature maps (Jogin et al., 2018). This operation is followed by a batch normalization, as the name indicates, is a simple normalization of the feature maps to be of the same order to facilitate learning of the network. Then, the ReLu function is applied to introduce non-linearity in the process. At the end of these operations, max-Pooling is used to reduce the size of the maps without losing the main information. This loss of resolution will be compensated in the decoder part. Each decoder starts with an up-sampling layer that is used to increase the resolution of the maps extracted in the encoder block. This up-sampling results from a transfer of the max-pooling indices from the encoder block which drastically reduces the computation time compared to other encoder-decoder structures such as Unet (Siam et al., 2018). As in the encoder, convolution, batch normalization and ReLu operations are performed sequentially in the decoder to form the network. At the end of all these operations, we added a convolution and a regression layer.

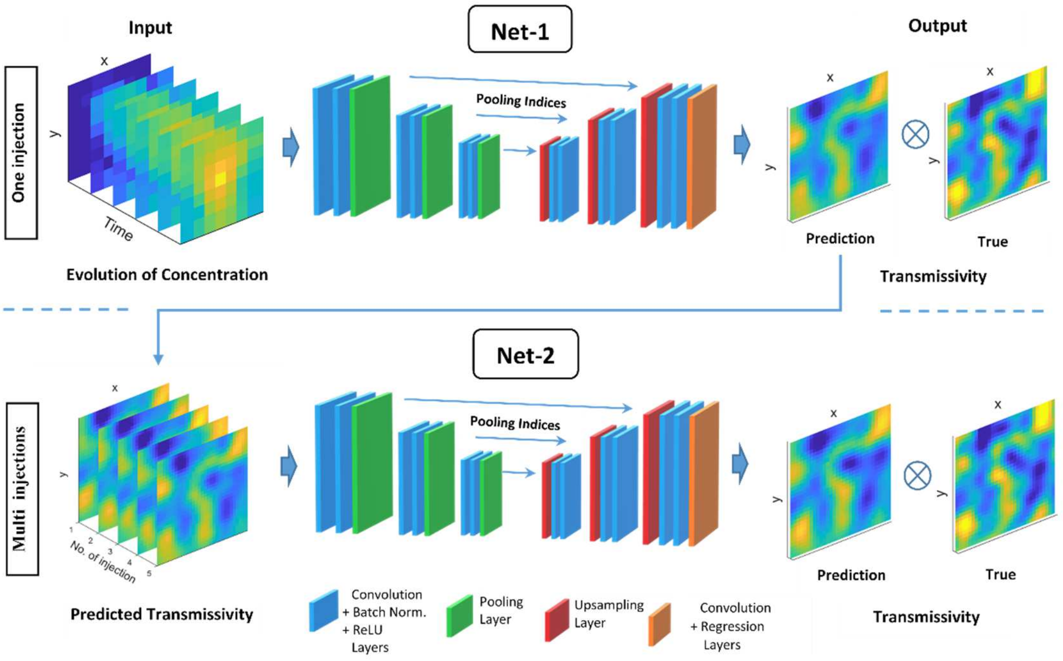

Figure 1: Graphical representation of the SegNet network modified to reconstruct the transmissivity field from concentration data for different injection configurations, named as CNN-2T. The upper network (Net-1) is designed to interpret data from a single injection into a single well. The lower network (Net-2), on the other hand, performs a transmissivity prediction

from the combination of multiple tracer tests which reprocesses the output of the upper network from each individual injection.

In this study, we test two schemes to reconstruct the transmissivity field (**Figure 1**): one is designed to process data collected with a single injection (Net-1) while the other processes data from multiple injections performed separately (Net-2). The latter uses the transmissivity fields obtained with a single injection as input to the second network to reconstruct a transmissivity model with multiple injections. Thus, the second network performs a joint inversion of the data to improve the prediction accuracy.

Both architectures are built with 3 encoders which relate to 3 decoders and the whole is formed of 45 layers. The architectures are trained in the same way, but the nature of the input data is totally different. The first one uses the concentration data measured in the boreholes at different time steps. Then, these punctual observations are spatially interpolated to bring this data to the same size as the transmissivity field. At the end, for each time step, an image is obtained to construct the spatio-temporal concentration data as a matrix with 3D. However, in the second network, the input data are the transmissivity fields obtained with the first network derived from the single injection interpretation. Thus, in the second network, the input and output have the same nature; this network provides a correction of the transmissivity field obtained by single injection.

3. Application

3.1 Construction of training data

This section is dedicated to the construction of the training dataset to train the Segnet network. To accomplish, we use a geostatistical code to generate 25,000 transmissivity fields in which the distribution of $\log_{10}T$ is constructed randomly using a Gaussian variogram. Variogram properties vary 0.1-2.5 for sill, 3-5 for mean, and 3-15 m for range. The values of the generated transmissivity models range from 10^{-6} to 10^{0} m^2/s in six orders of magnitude.

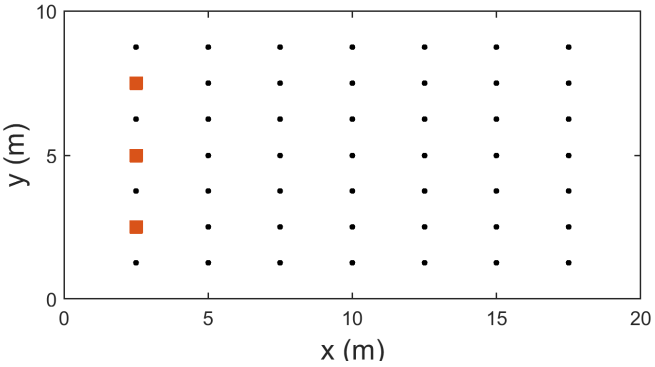

Figure 2: The tracing test is realized with 49 observation (dots) and 3 injection (red squares) wells in a studied area of 20 m × 10 m. Background flow is from left to right.

We assign these models to a confined aquifer with dimensions of (20m × 10m) and 5m for thickness (**Figure 2**). On this aquifer, we have set up 49 wells, 3 of which are dedicated to the injection of the tracer and the others to the monitoring of the evolution in time of the tracer concentration. Then, the set of transmissivity fields is used in the forward problem described in the previous sections to numerically calculate the concentration data associated to the tracer tests conducted sequentially on the three wells. The tracer is injected with a constant flow rate of 50m^3/day and a concentration of 50 mol/m^3. Each tracer test lasts 15 days. We also impose a hydraulic gradient of 0.1% between the right and left boundaries of the medium to simulate a regional flow. For simplicity, we assign constant values to the porosity (ϕ = 0.3), and longitudinal and transverse dispersivities of the aquifer (2.0 m and 0.2 m).

The concentration data acquired on the 49 boreholes over a period of 15 days with a time step of 0.5 days are transformed with this formulation to avoid strong contrasts between the data:

$$\hat{c} = \log_{10}(c + 1). \tag{11}$$

Then these data are spatially interpolated with nearest method at each time step to make the concentration data the same size as the transmissivity models. At the end, the input data are assembled as a 3D matrix (32 × 32 × 30). The same operation is performed for each hydraulic transmissivity model and each tracer test. Finally, we will try in the training phase to link the concentration data image with multiple channels (32 × 32 × 30) to another image of transmissivity with the same resolution (32 × 32).

For the training and testing phases, we separated 25,000 transmissivity fields with corresponding concentration data into 3 sets: 18,000 models for the training and 2,000 models for the validation. The remaining 5000 models are used in the test phase to evaluate the efficiency of the network in handling unseen models. All these data are organized according to the number of injections and take in total 2 GB.

3.2 Learning and validation of the neural network

As described in Section 2.2, we build two architectures in this study, Net-1 and Net-2, to process the single-injection test and the multi-injection test respectively. For the first network (Net-1), it will be trained three times, one for each injection tests. The trainings of both networks are performed with the ADAM optimization algorithm implemented in MATLAB on a Dell Precision Tower 5810, single GPU NVIDIA Quadro K2200. For Net-1, the training is performed with 60 epochs in 82 minutes, with a learning rate initiated at 0.01 and decreasing by 0.1 for every 30 epochs. However, the second network Net-2 is trained in 35 minutes, using a constant learning rate of 0.01 with 30 epochs. The number of epochs in both training is set manually to avoid the overfitting problem. The training quality of both networks is reported in **Figure 3**, where we presented the root mean square error between the predicted and true transmissivity models used in the training and validation.

To quantitatively assess the predictions accuracy of the networks on the unseen data, we use the coefficient of determination (R^2) and the root mean square error (RMSE) as defined below

$$R^2 = 1 - \frac{SSR}{SST}, \tag{12}$$

$$RMSE = \sqrt{\sum_1^n \frac{\left(s_{pred} - s_{true}\right)^2}{n}}, \tag{13}$$

where $SSR = \sum_1^n \left(s_{pred} - s_{true}\right)^2$, $SST = \sum_1^n (s_{true} - \bar{s}_{true})^2$ with s_{true} and s_{pred} denote the true and predicted logarithm of transmissivity, $s = -log_{10}T$, respectively; $\bar{s}_{true}$ is the mean of s_{true} and n is the number of pixels (32 × 32 in this case).

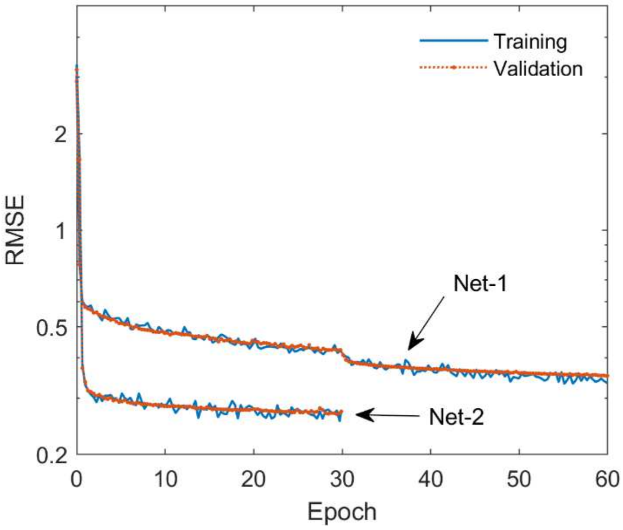

Figure 3: Loss function expressed in mean square error to track the quality of the learning process performed with 18,000 samples. The learning process of Net-1 required 60 epochs to reach a final convergence, while the learning of the second network (Net-2) required only 30 epochs. Net-2 uses all the predicted transmissivities fields from the Net-1 for each individual injection test to effectively improve the accuracy of predictions

To illustrate and discuss the quality of the inversions obtained with both networks, we have graphically represented the predictions of three models chosen among the test models. These models have different degree of heterogeneity, which is very high in model 3, moderate in the second and very low in the first model (**Figure 4**). This difference in heterogeneity is due to the parameters of the variogram used in generation, which were chosen randomly. The present discussion concerns the Net-1 trained with single injection data realized at the second well with coordinates (2.5 m, 5 m). The metric evaluations of the application of Net-1 on 5000 models in the test set shows that the accuracy of predictions depends on the complexity of the target model with the averages of determination coefficient and root mean square $R^2 = 0.68$ and RMSE = 0.36 respectively. This is also confirmed by the prediction reconstructions of three models presented in the figure 4 where the model with simple heterogeneity is better reconstructed (Exp-1, $R^2 = 0.88$) compared to the models with moderate (Exp-2, $R^2 = 0.56$) and complex (Exp-3, $R^2 = 0.42$) heterogeneities. We also observe that these reconstructions are better in areas close to the injection well. This is normal since in the areas far from the injection wells, the concentration decreases and the acquisition time stops at 15 days, which makes it difficult to obtain information on the hydraulic properties.

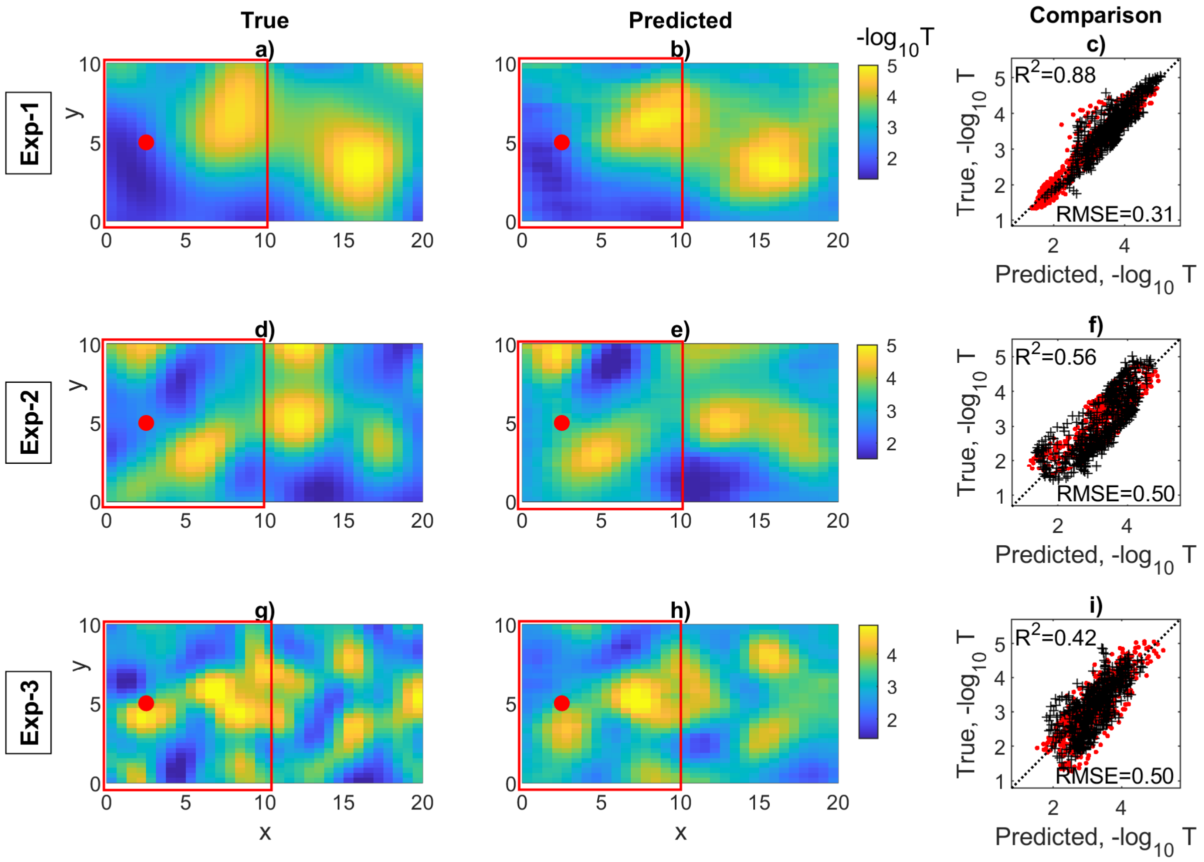

Figure 4: Predictions using Net-1 with data obtained from a single injection (at the red point) - Case C0. The accuracy of these predictions depends on the degree of heterogeneity of the models in which the model with simple heterogeneity is better reconstructed than the models with moderate and complex heterogeneities. Furthermore, the area near the injection well (delimited by a red box) is well mapped compared to the distant areas, as shown by the correlation of the transmissivity data of this area (red dots in Fig 4.cfi) with that of the true model.

In order to improve the quality of the predictions, we will combine in Net-2 all the predictions derived from Net-1 with the single injection. Indeed, the assimilation of three transmissivity fields coming from three injections to form the Net-2 should allow a better reconstruction of the whole area. On the other hand, another configuration is also feasible which consists in combining all the data from the three injections to train the network at the same time. But this approach obviously requires more memory because the concentration data will be voluminous (32×32×30×3) which makes the training very complex. On the other hand, by separating the data of each injection, we reduce the memory, and we can analyze the information brought by each tracer test.

The training of Net-2 is relatively fast compared to the first network because the training data size is reduced to 3-channels (32×32×3). It is also very easy to relate the predicted transmissivity to the real transmissivity fields, which have many similarities. The performance test of this network was also performed on 5000 unseen test models. The prediction results of all these models have an average of $R^2 = 0.83$ and RMSE = 0.2. The details of the distribution of R^2 and RMSE are summarized in **Figure 10** with a comparison to the previous Net-1 results. Based on this metric evaluation, this network seems to correctly reproduce the aquifer heterogeneities. This is also confirmed by the analysis of the predictions of the three models which illustrates that with Net-2, the reconstructions were well improved for all models compared to the results with first Net-1 with single tracer test (see **Figure 5**).

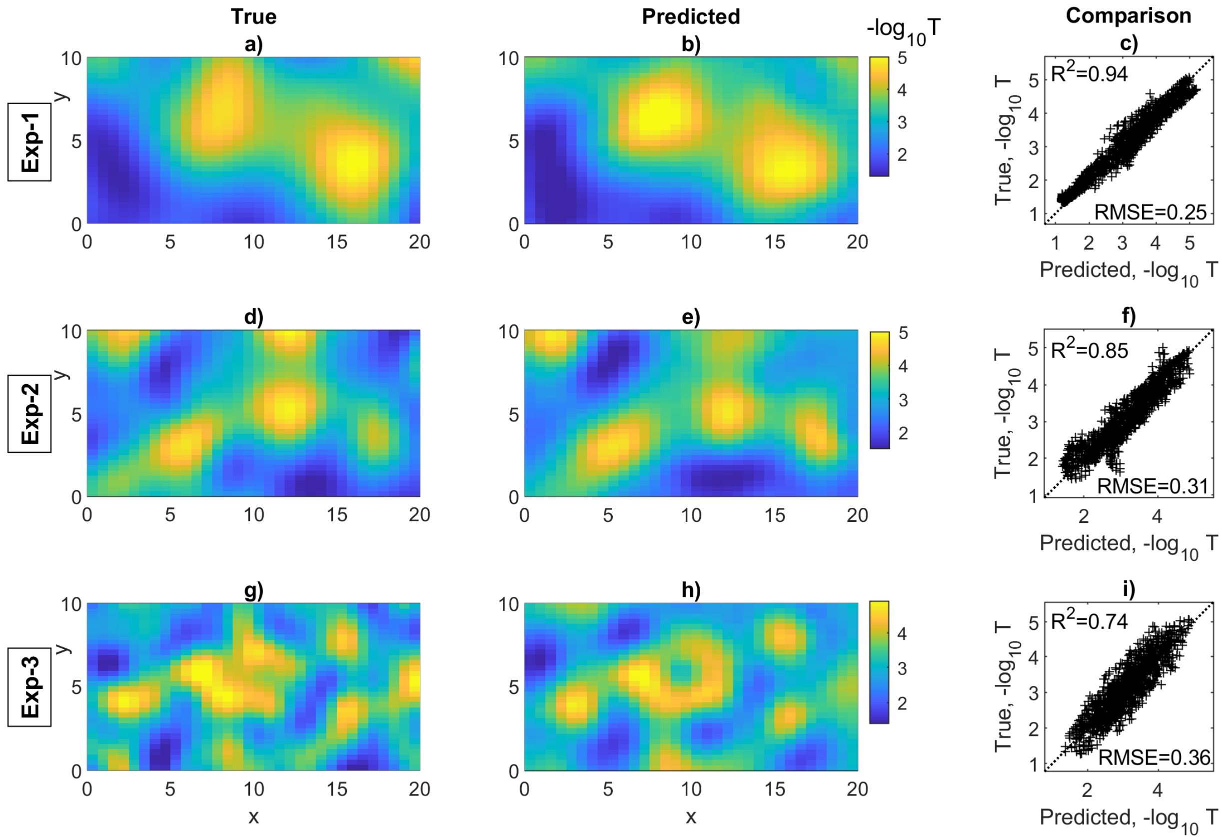

Figure 5: Predictions using Net-2 where the predicted transmissivity fields from Net-1 are used as input. This combination of data from three tracer tests reconstructs the aquifers heterogeneities accurately and leads to an improvement in the predictions over the Net-1 results in Figure 4.

3.3 Effect of the resolution of the observation data

In this section, we will investigate how changing the quantity of piezometers and the duration of concentration data sampling affects the quality of predictions.

a. Number of measurement wells (in space)

In this test, we only reduce the number of wells used to track the concentration changes to 25 wells; in other words, we remove almost half of the wells from the initial configuration used in the previous cases (**Figure 6**). However, we keep the same sampling duration used in the previous cases. Then the data acquired with this new configuration is used to train the networks without modifying their structures. The trained network Net-2 was also used to predict unseen test models and the results obtained have average values of $R^2 = 0.70$ and RMSE = 0.33.

This metric evaluation illustrates a degradation in the reconstruction accuracy compared to the result of the previous case with a dense distribution of wells. However, the concentration data are still largely sufficient to provide the main features of the models. As confirmed by the predictions of the reference models in which it is found that the reconstructions of the models with simple and moderate heterogeneities remain satisfactory (**Figure 7**). However, the third model with complex heterogeneity the reconstruction obtained is very smooth, as the lack of data did not allow to perfectly identify the heterogeneities of small sizes.

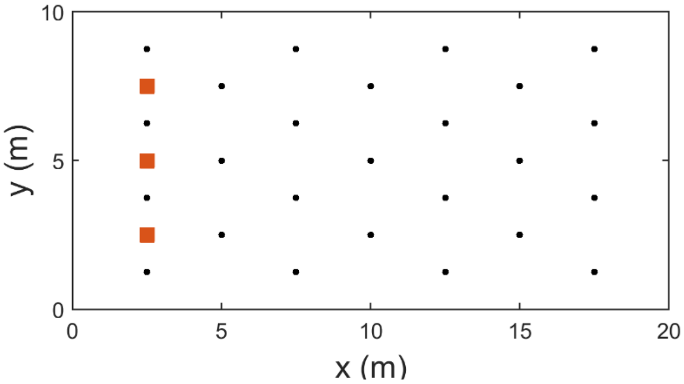

Figure 6: A novel setup for measuring concentration data using only 25 wells (black dots) and three injections (red squares) to study the effect of lack of data on prediction quality.

We then reduced the number of observation wells to 16, arranged in a regular array of (5.0 m × 2.5 m). Analysis of the prediction accuracy with this poor configuration reveals a significant degradation in the quality of the reconstructions with average of $R^2 = 0.57$ and RMSE = 0.38. All the information on the metric evaluation of the case is reported in

Table 1. We conclude that the accuracy of inversion results with deep learning methods like other classical inversions with deterministic (Gauss-Newton) or stochastic (MCMC) methods depends strongly on the number of wells used to collect the data. Thus, a configuration with a very limited number of piezometers will not be able to capture all the characteristics of a highly heterogeneous aquifer but will provide a too smooth mapping of the hydraulic transmissivity.

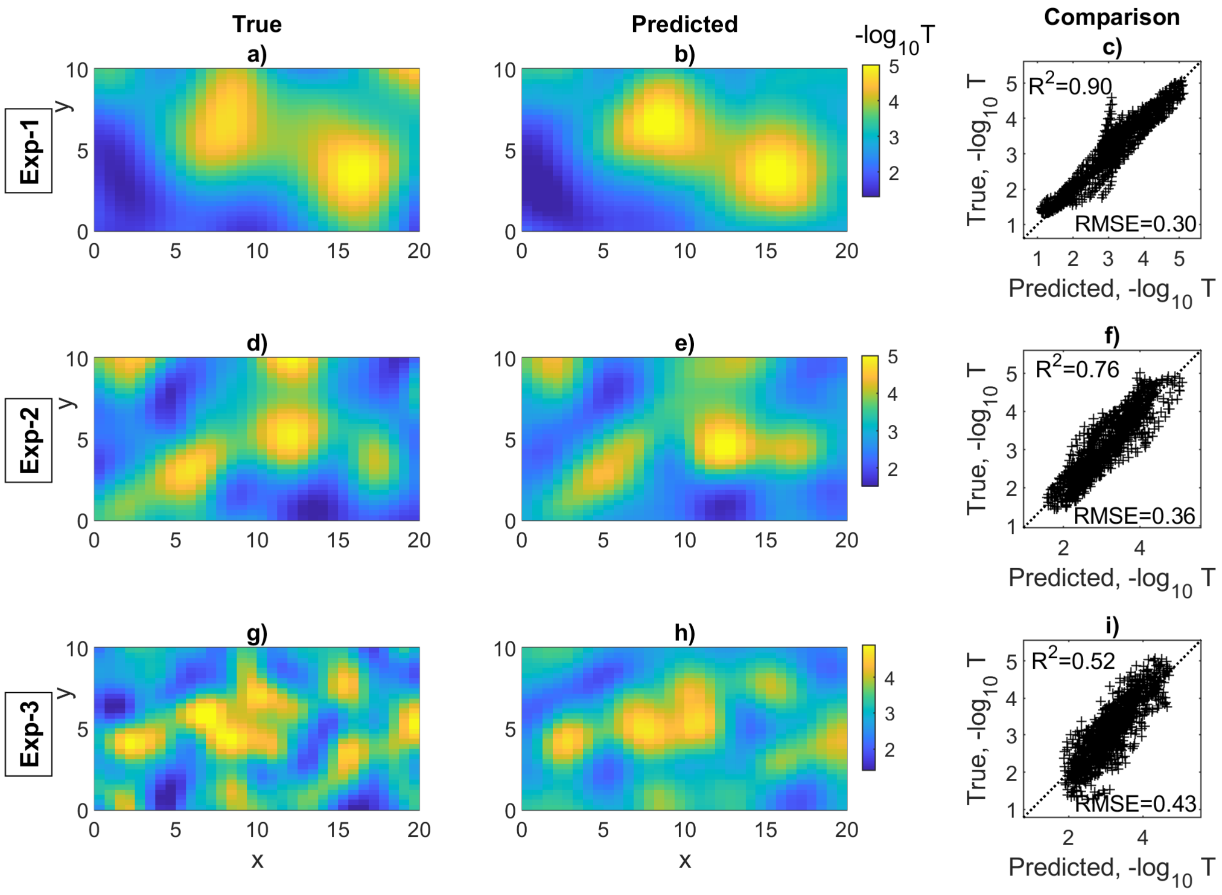

Figure 7: Predictions of the three models from concentration data recorded at 25 wells (Case C2). In general, data reduction leads to a deterioration in the quality of predictions, especially in the case of very heterogeneous models.

b. Number of measurements in time

In this section, we analyze the impact of reducing the duration of tracer concentration monitoring from 15 days to 5 and 3 days on the predictions. In practice, we only shortened the number of training data channels used in the construction of Net-1 with the first configuration (with 49 piezometers). Thus, the data that will be taken in this analysis will have these sizes (32×32×10) and (32×32×6) for the duration of 5 and 3 days respectively. We re-train the networks with these new data and use them to predict the test models. The results obtained for the two acquisition times of 5 and 3 days in **Table 1** have these averages (R^2 = 0.81, RMSE = 0.30) and (R^2 = 0.74, RMSE = 0.33) respectively. From these evaluations, there is a net decrease in prediction quality with shortening of the acquisition time. The predictions of the reference models confirm this trend but also show that this degradation occurs mainly on the areas far from the injection zone (**Figure 8**). Indeed, the use of a short acquisition time does not allow collecting enough information on the distant areas where the tracer reaches late. The accuracy of the reconstruction with the deep learning algorithm depends on the spatial coverage of the piezometer and the temporal acquisition scheme used to track changes in tracer concentration over time. Dense coverage with a long acquisition time allows for better identification of transmissivity heterogeneities.

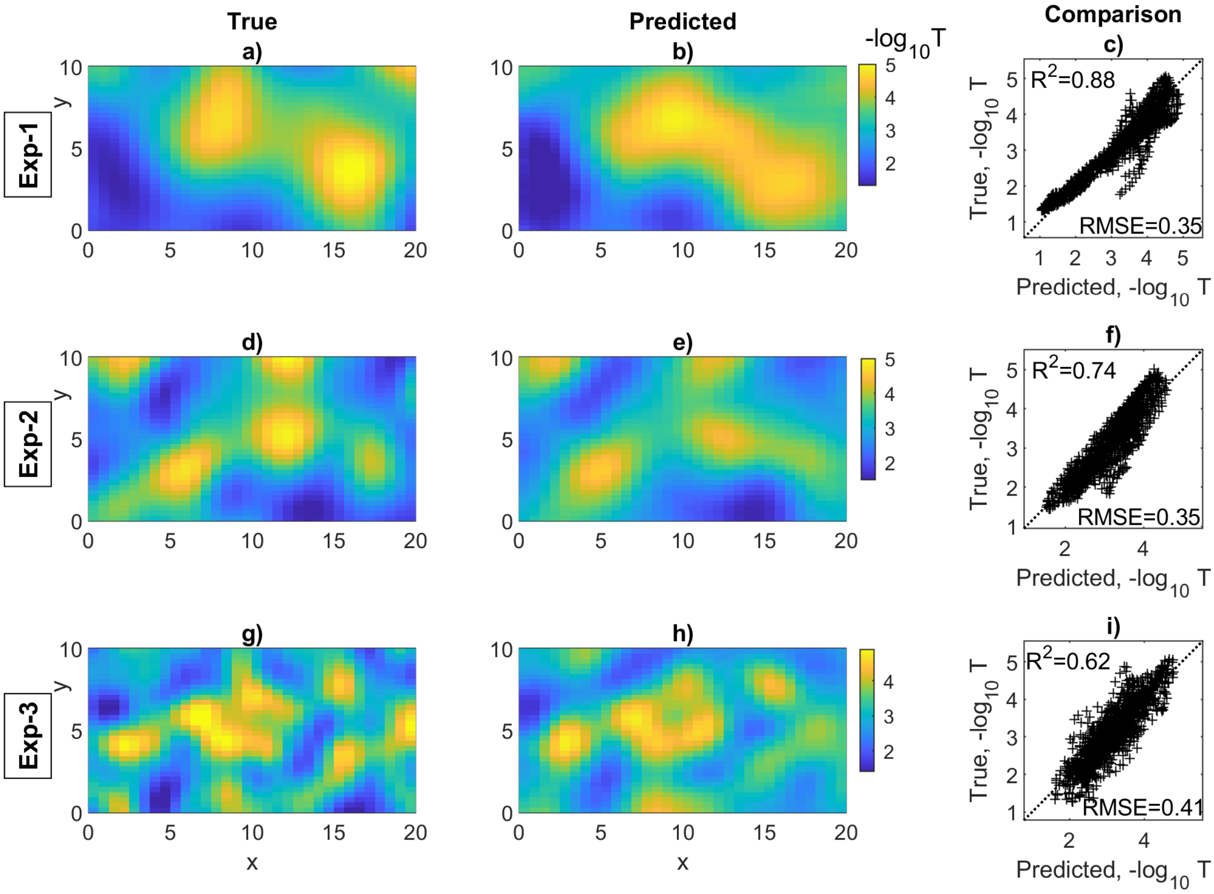

Figure 8: Predictions obtained from concentration data collected only over 3 days (case C3). The shortening of the acquisition time affects the quality of predictions of regions far away (x > 10 m) from the injected wells.

3.4 Effect of observation uncertainty

Most often in tracer tests, the recorded data may carry a noise signal that may affect the inversion result. In this section, we analyze the noise impact on the result of the deep learning algorithm. To do so, we contaminate the concentration data used in the test models with a Gaussian noise of 15% and 25% and check the result of applying the network already built in the previous section on these contaminated data. The result of this application is presented in the table, where the averages of the correlation coefficients (R^2 = 0.79 for 15% noise, and R^2 = 0.76 for 25% noise) indicate a slight degradation of the noise quality with the increase of noise magnitude.

The transmissivity predictions of the reference models also show that the 25% noise does not prevent efficient reconstructions of hydraulic transmissivity fields (**Figure 9**). This minimal effect of noise has also been noted in applications of CNN techniques on the inversion of electrical (Vu & Jardani, 2021), seismic data (Wu & Lin, 2018; Wei & Chen, 2019). This could be explained by the fact that the data in CNN networks are not processed individually but in spatial connection with each other and in this case the noise does not have a significant impact on the alteration of the information carried by the input image.

To clarify the influence of each interference in the prediction accuracy, details of the distribution of R^2 and RMSE are summarized in **Table 1** for three examples and the average of 5000 testing models with its histogram illustrated in **Figure 10**.

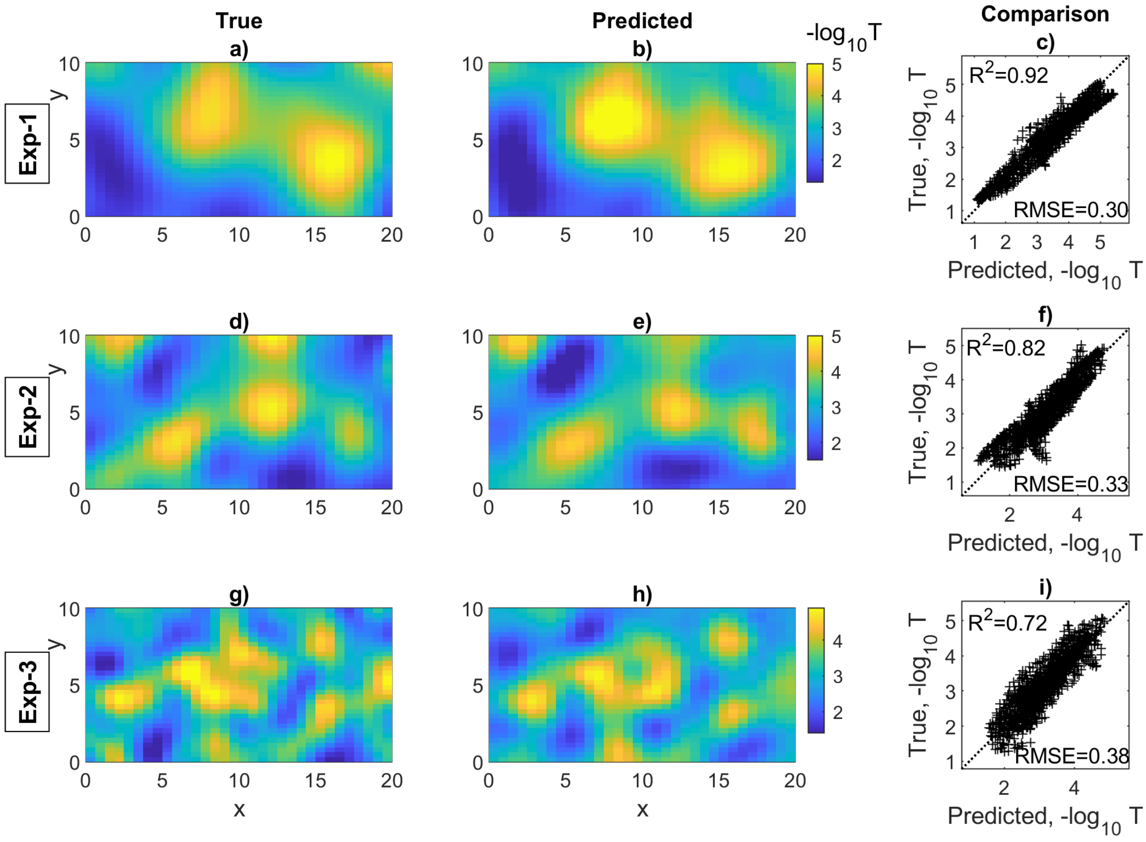

Figure 9: Inversion results from concentration data contaminated with 25% Gaussian noise (Case C4). These predictions show that noise has a very limited effect on the quality of the predictions since the CNN concept relies on the spatial analysis of the whole data set rather than on individual data points.

Table 1: Summary of metric evaluations R^2 and RMSE of the 3 reference models and the 5000 test models (shown as $\frac{R^2}{\mathrm{RMSE}}$).

Data type	*One injection*	*Full data*	*Lower resolution*		*Shorter obs.*		*With noise*	
			25 wells	16 wells	3 days	5 days	25%	15%
	(C0)	(C1)	(C2)		(C3)		(C4)	
Exp-1	$\frac{\mathbf{0.88}}{0.31}$	$\frac{\mathbf{0.94}}{0.25}$	$\frac{\mathbf{0.90}}{0.30}$	$\frac{\mathbf{0.80}}{0.37}$	$\frac{\mathbf{0.88}}{0.35}$	$\frac{\mathbf{0.90}}{0.30}$	$\frac{\mathbf{0.92}}{0.30}$	$\frac{\mathbf{0.93}}{0.27}$
Exp-2	$\frac{\mathbf{0.56}}{0.50}$	$\frac{\mathbf{0.85}}{0.31}$	$\frac{\mathbf{0.76}}{0.36}$	$\frac{\mathbf{0.68}}{0.38}$	$\frac{\mathbf{0.74}}{0.35}$	$\frac{\mathbf{0.86}}{0.29}$	$\frac{\mathbf{0.82}}{0.33}$	$\frac{\mathbf{0.83}}{0.33}$
Exp-3	$\frac{\mathbf{0.42}}{0.50}$	$\frac{\mathbf{0.74}}{0.36}$	$\frac{\mathbf{0.52}}{0.43}$	$\frac{\mathbf{0.17}}{0.55}$	$\frac{\mathbf{0.62}}{0.41}$	$\frac{\mathbf{0.75}}{0.37}$	$\frac{\mathbf{0.72}}{0.38}$	$\frac{\mathbf{0.73}}{0.36}$
Average of 5000 samples	$\frac{\mathbf{0.68}}{0.36}$	$\frac{\mathbf{0.83}}{0.28}$	$\frac{\mathbf{0.70}}{0.33}$	$\frac{\mathbf{0.57}}{0.38}$	$\frac{\mathbf{0.74}}{0.33}$	$\frac{\mathbf{0.81}}{0.30}$	$\frac{\mathbf{0.76}}{0.33}$	$\frac{\mathbf{0.79}}{0.31}$

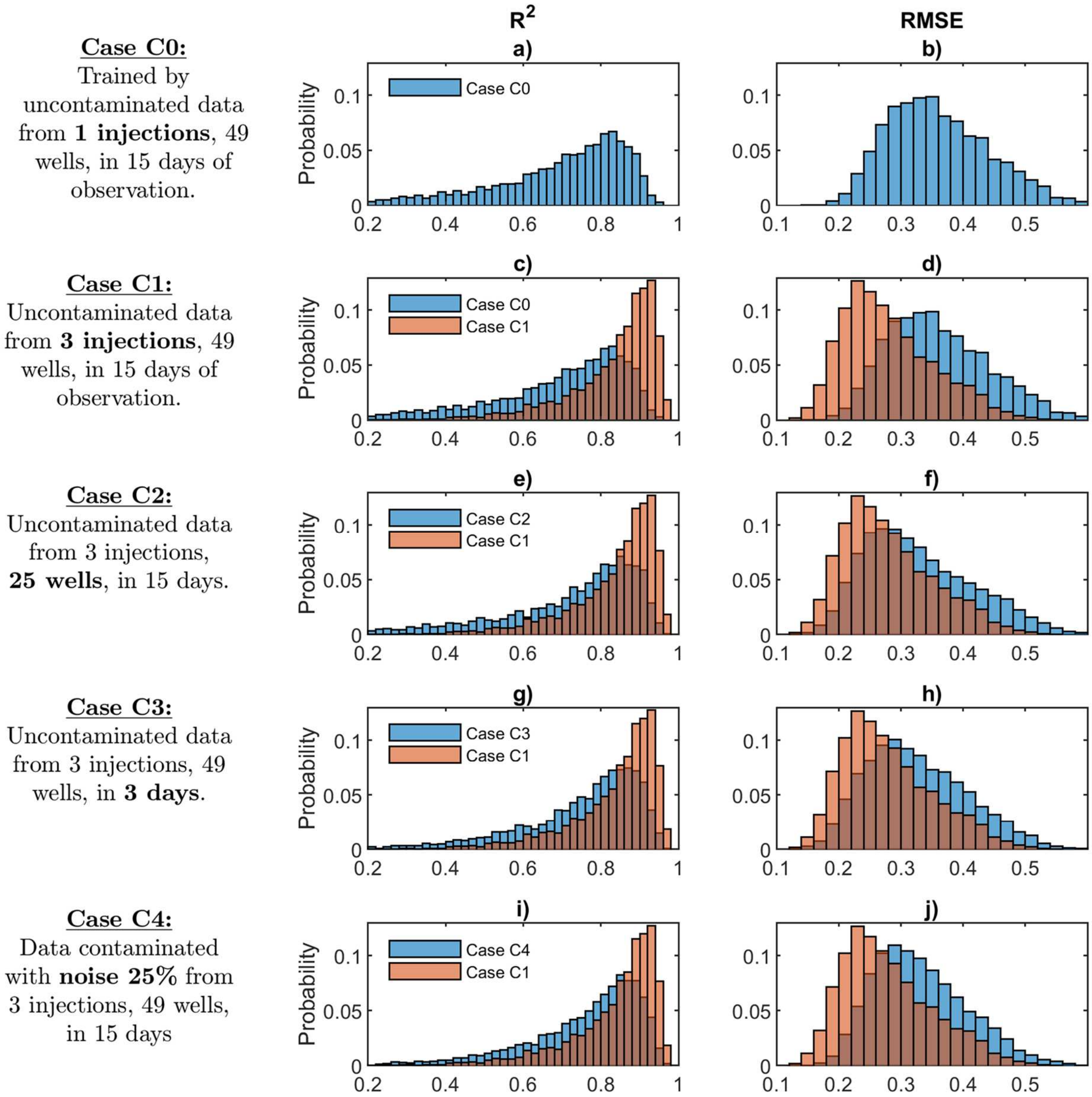

Figure 10: Distributions of the metric parameters R^2 and RMSE determined from the comparison between predicted and real models and used to evaluate the accuracy of predictions. We recall that these evaluations were made on 5000 models.

4. Influence of the features of training models

In the inverse problem, it is known that the solution is not unique and to reduce the number of solutions, it is necessary to include prior information in the optimization procedure. In the deep learning algorithms, prior information is also included through the choice of the characteristics of generated transmissivity fields of training data. To evaluate how the nature of the training models can influence the inversion results, we apply the previous network trained with a Gaussian variogram on the inversion of the concentration data derived from the non-Gaussian transmissivity models. We generated three models characterized by abrupt changes in transmissivity values that are completely different from the smooth models used in the network training. The predictions resulting from these applications have been reported in **Figure 11** with comparisons between the predicted and true models. Based on these comparisons, we can clearly notice that the network is able to provide the main heterogeneities of the first and second models. But in detail, these models show many anomalies related to the presence of variabilities in transmissivity on the areas where the values are homogeneous. This is related to the influence of the nature of the training models. However, on the third model with complex heterogeneity, the network fails to provide a good reconstruction of the field. These results are normal because deep learning algorithms only approximate an inversion operator that is highly dependent on the nature of the training data and thus cannot be used to predict models with unlearned features. Thus, the choice of the feature of the training models is a crucial step in the deep learning algorithm and this choice must be based on the prior information collected on the aquifer under study.

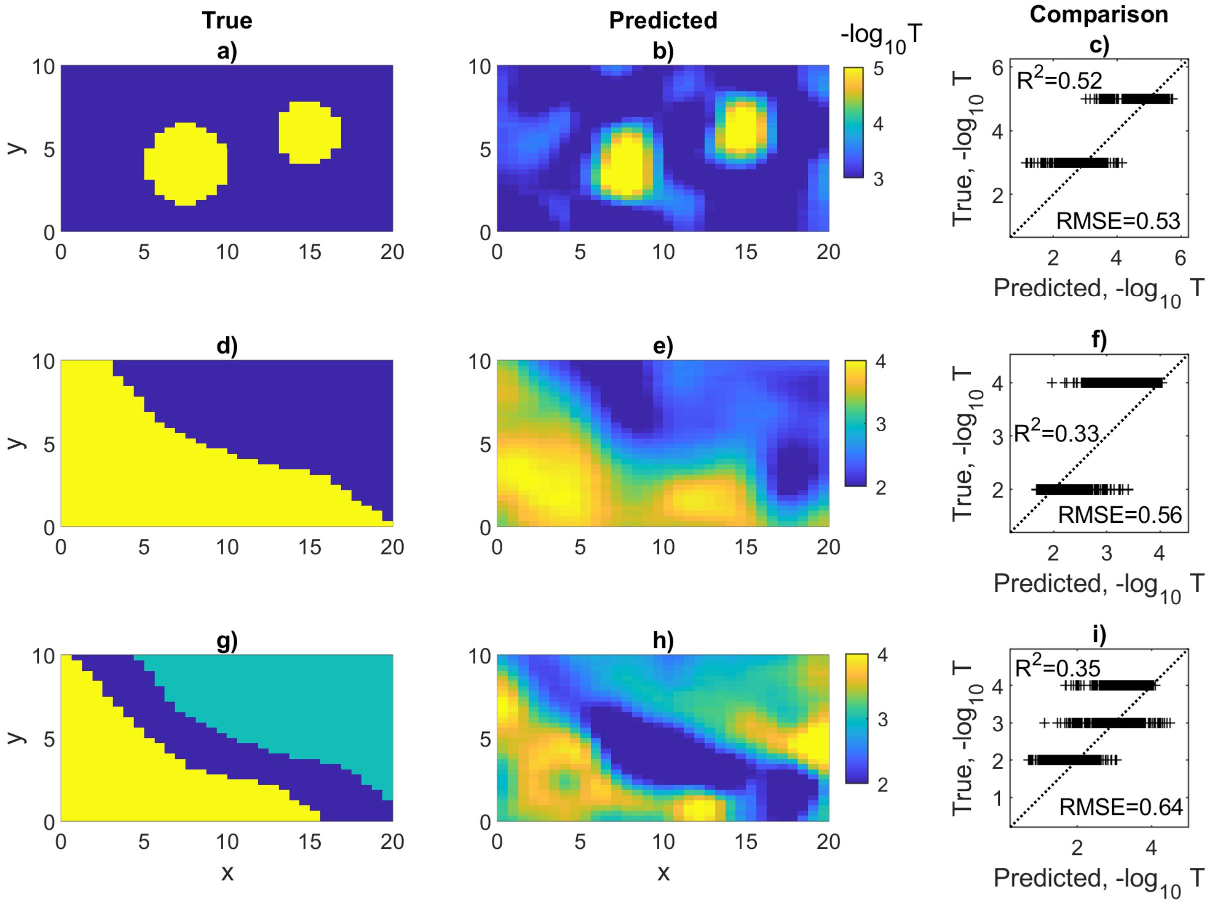

Figure 11: Applying a trained network with transmissivity models generated with a Gaussian variogram to predict transmissivity models with discontinuous heterogeneities. Qualitatively, the predictions reproduce the main transmissivity distributions found in the real fields, especially in the first and second models. However, in detail, they fail to provide accurate reconstructions. Therefore, inversion with the deep learning algorithm cannot predict models with features not seen in the learning steps.

5. Comparison of CNN-2T with conventional Gauss-Newton algorithm

We devote this section to comparing the results obtained with our CNN-2T deep learning algorithm and the Gauss-Newton code. Gauss-Newton is a popular deterministic algorithm that searches for the best transmissivity model that can fit the concentration data by minimizing an objective function formulated as follows:

$$L = (f(\boldsymbol{m}) - \hat{\boldsymbol{c}}_{obs})^T\, (f(\boldsymbol{m}) - \hat{\boldsymbol{c}}_{obs}) + \alpha\, (\boldsymbol{m} - \boldsymbol{m}^*)^T\, \boldsymbol{G}^{-1}\, (\boldsymbol{m} - \boldsymbol{m}^*), \tag{14}$$

where $\boldsymbol{m}$ is the logarithm of the transmissivity field to be determined from the minimization of this objective function. $\hat{\boldsymbol{c}}_{obs}$ denotes the measured concentration vector, $f(\boldsymbol{m})$ is the forward problem used to compute numerically the concentration vector corresponding to the transmissivity model $\boldsymbol{m}$. $\boldsymbol{G}$ is the covariance matrix used with the mean of field $\boldsymbol{m}^*$ as prior information to constraint the inversion. α is a parameter of regularization. The minimization of the objective function relies on this iterative formulation:

$$\boldsymbol{m}^{k+1} = \boldsymbol{m}^k + \beta \Delta \boldsymbol{m}, \tag{15}$$

where

$$\Delta \boldsymbol{m} = [\boldsymbol{J}^T \boldsymbol{J} + \alpha\, \boldsymbol{G}^{-1}]^{-1} \boldsymbol{J}^T \left(\hat{\boldsymbol{c}}_{obs} - f(\boldsymbol{m})\right), \tag{16}$$

with coefficient β is a scalar determined from minimizing objective function *L* using a line search code; ***J*** is a Jacobian matrix computed with finite difference:

$$\boldsymbol{J} = \frac{\partial f(\boldsymbol{m})}{\partial \boldsymbol{m}}, \tag{17}$$

The calculation of the Jacobian matrix requires a very repetitive and heavy computation of the forward problem, in this case, for each iteration, the direct problem is solved (1024 = 32×32) times. The Jacobian could be computed with the adjoint state method, which is faster than the finite difference method, but it is complex to implement (Vu et al, 2019; Sun & Yeh, 1990). The process is repeated until a local minimum is reached that minimizes the objective function. This algorithm was applied on the three models used as reference in the previous sections and the inversion of each model took 3 days using a CPU (Intel Xeon 3.5GHz, 6 cores, 12 threads). The results of these inversions faithfully reproduce the main characteristics of the real models with the same degree of accuracy as the CNN-2T algorithm. Each inference of Gauss-Newton inversion method consumes the same time with building entire process of CNN-2T, including time for dataset generation and training network, while CNN-2T performs each inference only on few milliseconds.

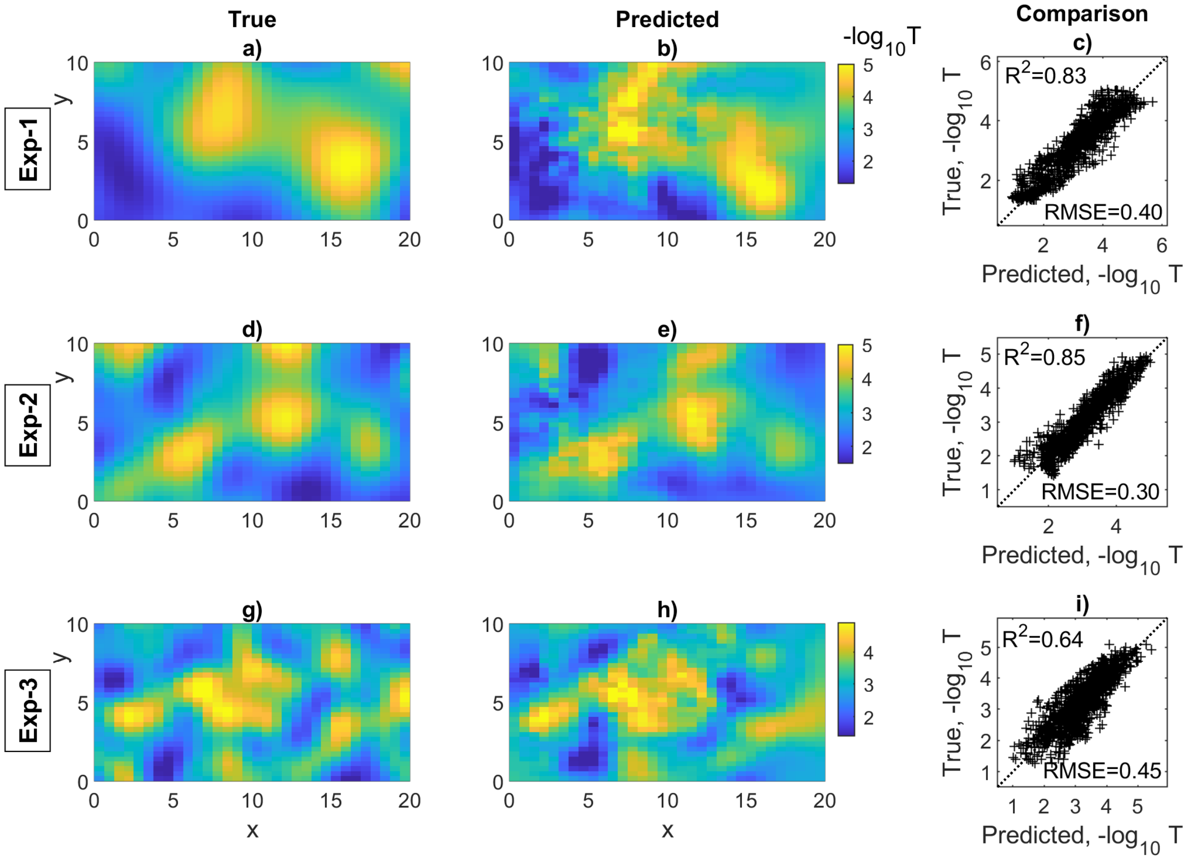

Figure 12: Predictions from a conventional inverted method of which the accuracy is of the same order with ones from the Neural Network in Figure 5. A conventional method takes days to process each inversion comparing to milliseconds by a neural network approach.

As explained in the work of Jardani et al. 2022, deterministic inversion methods share with the inversion method their strong dependence on the a priori model. In deterministic algorithms, the a priori model is used as a regularization term of the objective function, so it must have a derivable form. On the other hand, in machine learning, the a priori model is used to construct training data without any condition on its derivability. It also shares their dependence on the amount of measurement data and that the quality of the prediction improves in both methods when the piezometric information are sufficient to capture all the heterogeneities of the studied aquifer. Regarding the impact of measurement noise on the results, machine learning techniques are slightly less impacted compared to deterministic methods. Because in machine learning, the data processing is performed on the spatial analysis of several information instead of individual data as it is the case in deterministic algorithms. The results of the Gauss-Newton inversion are local minima, so they depend on the choice of the initial model. In contrast, machine learning provides a global solution. The machine learning method does not require any linearization of the forward problem, but it does require a large amount of forward problem computation to construct the training data. The time spent generating this training dataset can be the same or less than that needed to find a local minimum with Gauss-Newton code using the finite difference method. The reason it will take less time is that training an inversion network with simple and moderate heterogeneities can be done with a very limited data set. We recall that in our case we used a variogram with randomly chosen parameters which leads to create a large variability between models, which makes the training is done with a large volume of training data.

Conclusion

In this work, we propose a novel approach that inverts tracer test data to reconstruct the spatial variability of the hydraulic transmissivity field. The approach is based on a deep learning neural network to approach the inversion function. The proposed neural network architecture - Segnet - uses the concept of convolutional neural networks with encoder-decoder structure, which have shown significant success in image processing in recent years. The Segnet network was originally set up for segmentation tasks, but in this paper, it has been modified to perform a regression task in order to relate the temporal concentration data to the transmissivity field. The inversion processing was performed sequentially in two steps and via two connected networks with the same structure: The first network was built to link the concentration data recorded with a single tracer test and the transmissivity field. Then, this network is trained with data of each tracer test performed separately.

The transmissivity fields determined from the individual tracer data are then incorporated into the second network as input data, which is linked to the real transmissivity field to improve the reconstructions. This second network performs a joint inversion of data from multiple tracer tests and corrects the transmissivity fields determined separately from each individual tracer test.

The networks were trained from a large synthetic database where the transmissivity fields were generated geostatistically and the concentration data by the coupled numerical solutions of the groundwater flow and transport equations. Several sensitivity analyses were conducted on these networks to evaluate the prediction accuracy, which was also compared with that achieved by a deterministic Gauss-Newton method. Based on these results, we conclude that:

The network succeeded in accurately reconstructing the heterogeneities of the transmissivity field. The reconstruction accuracy though shows a clear dependence on the scale of training datasets, whose generation is the hardest task in a deep learning approach. However, the time needed to generate datasets is of the same magnitude as that required in solving a local minimum in a Gauss-Newton algorithm which computes the Jacobian matrix using the finite difference method. In addition, the size of the training datasets to achieve a perfect approximation of the inverse function depends strongly on the nature of the heterogeneity of the target field. Learning can be achieved with a small number of training models if the generated models have a low degree of heterogeneity. Recall that in this work, the models are all generated according to a random scheme in which the variogram parameters are also randomly selected to ensure the generality in the generation. As a result, a large number of training models are required to reconstruct an accurate prediction.

Furthermore, the quality of the neural network predictions relies on the size of the concentration data, i.e., the number of wells and the duration chosen for tracer migration monitoring. Accuracy improves as monitoring data are sufficient to cover the main heterogeneities in the target field. This dependence on data coverage is also observed in both classical deterministic and stochastic inversion methods. Concerning the effect of concentration data noise on forecasts, the tests revealed that it had a marginal impact on the network performance. This is due to the spatial analysis performed by the convolutional process, in which the data are related regionally rather than being handled separately as in classical inversion methods.

We believe that the new generation of deep learning algorithms based on the concept of convolutional neural networks are effective tools for mapping hydraulic properties from tracing or pumping test data. These tools deserve to be explored in hydraulic and geophysical tomography and to show all their aspects, we plan in our next work to apply them on a real field data.

References

Akin S. (2005). Tracer model identification using artificial neural networks. *Water Resources Research*, 41(10).

Apolinario M., Huaman Bustamante S., Morales G. & Diaz D. (2019). Estimation of 2D Velocity Model using Acoustic Signals and Convolutional Neural Networks. *2019 IEEE XXVI International Conference on Electronics, Electrical Engineering and Computing (INTERCON).*

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.

Bailey R. & Baù D. (2012). Estimating geostatistical parameters and spatially-variable hydraulic conductivity within a catchment system using an ensemble smoother. *Hydrol. Earth Syst. Sci.*, 16(2), 1607-7938.

Balkhair, K. S. (2002). Aquifer parameters determination for large diameter wells using neural network approach. Journal of Hydrology, 265(1), 118-128.

Bao J., Li L. & Redoloza F. (2020). Coupling ensemble smoother and deep learning with generative adversarial networks to deal with non-Gaussianity in flow and transport data assimilation. *Journal of Hydrology*, 590, 125443.

Bear, J. (1972). Dynamics of Fluids in Porous Media. American Elsevier.

Cardiff M., Barrash W. & Kitanidis P. K. (2013). Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. *Water Resources Research*, 49(11), 7311-7326.

Carrera J., Alcolea A., Medina A., Hidalgo J. & Slooten L. (2005). Inverse problem in hydrogeology. *Hydrogeology Journal*, 13(1), 206-222.

Cui T., Fox C. & O'Sullivan M. (2011). Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis Hastings algorithm. *Water Resources Research*, 47(10).

Datta-Gupta A., Lake L. & Pope G. (1995). Characterizing heterogeneous permeable media with spatial statistics and tracer data using sequential simulated annealing. *Mathematical Geology*, 27(6), 763-787.

Fahs H., Hayek M., Fahs M. & Younes A., 2014. An efficient numerical model for hydrodynamic parameterization in 2D fractured dual-porosity media. *Advances in Water Resources*, 63, 179-193.

Hassoun M. (1995). *Fundamentals of Artificial Neural Networks*. MIT Press.

Indolia S., Goswami A., Mishra S. & Asopa P. (2018). Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. *Procedia Computer Science*, 132, 679-688.

Jacob B., & Alexander C. H.-D. (2010). *Modeling Groundwater Flow and Contaminant Transport*. Springer Netherlands.

Jardani A., Revil A. & Dupont J. (2013). Stochastic joint inversion of hydrogeophysical data for salt tracer test monitoring and hydraulic conductivity imaging. *Advances in Water Resources*, 52, 62-77.

Jardani A., Dupont J., Revil A., Massei N., Fournier M. & Laignel B. (2012). Geostatistical inverse modeling of the transmissivity field of a heterogeneous alluvial aquifer under tidal influence. *Journal of Hydrology*, 472-473, 287-300.

Jardani A., Vu T. M. & Fischer P., 2022. Use of convolutional neural networks with encoder-decoder structure for predicting the inverse operator in hydraulic tomography. *Journal of Hydrology*, 604, 127233.

Jha M. K. & Datta B., 2011. Simulated annealing based simulation-optimization approach for identification of unknown contaminant sources in groundwater aquifers. *Desalination and Water Treatment*, 32(1-3), 79-85.

Jiménez S., Mariethoz G., Brauchler R. & Bayer P. (2016). Smart pilot points using reversible-jump Markov-chain Monte Carlo. *Water Resources Research*, 52(5), 3966-3983.

Jogin M., Mohana Madhulika M., Divya G., Meghana R. & Apoorva S. (2018). Feature Extraction using Convolution Neural Networks (CNN) and Deep Learning. *2018 3rd IEEE International Conference on Recent Trends in Electronics, Information Communication Technology (RTEICT)*, (pp. 2319-2323).

Kitanidis P. (1997). *Introduction to Geostatistics: Applications in Hydrogeology*. Cambridge University Press.

Kittilä A., Jalali M., Somogyvári M., Evans K., Saar M. & Kong X.-Z. (2020). Characterization of the effects of hydraulic stimulation with tracer-based temporal moment analysis and tomographic inversion. *Geothermics*, 86, 101820.

Laloy E., Hérault R., Jacques D. & Linde N. (2018). Training-Image Based Geostatistical Inversion Using a Spatial Generative Adversarial Neural Network. *Water Resources Research*, 54(1), 381-406.

LeCun Y. & Bengio Y. (1998). Convolutional Networks for Images, Speech, and Time Series. In *The Handbook of Brain Theory and Neural Networks* (pp. 255–258). Cambridge, MA, USA: MIT Press.

Lecun Y., Bottou L., Bengio Y. & Haffner P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.

Lee J. & Kitanidis P. (2014). Large-scale hydraulic tomography and joint inversion of head and tracer data using the Principal Component Geostatistical Approach (PCGA). *Water Resources Research*, 50(7), 5410-5427.

Lin G.-F. & Chen G.-R. (2006). An improved neural network approach to the determination of aquifer parameters. *Journal of Hydrology*, 316(1), 281-289.

Liu B., Guo Q., Li S., Liu B., Ren Y., Pang Y., Guo X., Liu L. & Jiang P. (2020). Deep Learning Inversion of Electrical Resistivity Data. *IEEE Transactions on Geoscience and Remote Sensing*, 58(8), 5715-5728.

Marçais J. and de Dreuzy J.-R. (2017). Prospective Interest of Deep Learning for Hydrological Inference. *Ground Water*, 55(8), 688-692.

O'Shea K. & Nash R. (2015). An Introduction to Convolutional Neural Networks. *arXiv*:1511.08458v2.

Puzyrev V. & Swidinsky A. (2019). Inversion of 1D frequency- and time-domain electromagnetic data with convolutional neural networks. *arXiv*:1912.00612.

Reuschen S., Xu T. & Nowak W. (2020). Bayesian inversion of hierarchical geostatistical models using a parallel-tempering sequential Gibbs MCMC. *Advances in Water Resources*, 141, 103614.

Saley A., Jardani A., Soueid Ahmed A., Antoine R., & Dupont J.-P. (2016). Hamiltonian Monte Carlo algorithm for the characterization of hydraulic conductivity from the heat tracing data. *Advances in Water Resources*, 97, 120-129.

Sanchez-León E., Leven C., Haslauer C. & Cirpka O. (2016). Combining 3D Hydraulic Tomography with Tracer Tests for Improved Transport Characterization. *Ground Water*, 54(4), 498-507.

Schwede R., Li W., Leven C. & Cirpka O. (2014). Three-dimensional geostatistical inversion of synthetic tomographic pumping and heat-tracer tests in a nested-cell setup. *Advances in Water Resources*, 63, 77-90.

Sen M., Datta-Gupta A., Stoffa P., Lake L. & Pope G. (1995). Stochastic Reservoir Modeling Using Simulated Annealing and Genetic Algorithm. *SPE Formation Evaluation*, 10(01), 49-56.

Siam M., Gamal M., Abdel-Razek M., Yogamani S., Jagersand M. & Zhang H. (2018). A Comparative Study of Real-Time Semantic Segmentation for Autonomous Driving. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 700-710.

Somogyvári M. & Bayer P. (2017). Field validation of thermal tracer tomography for reconstruction of aquifer heterogeneity. *Water Resources Research*, 53(6), 5070-5084.

Soueid Ahmed A., Zhou J., Jardani A., Revil A. & Dupont J. (2015). Image-guided inversion in steady-state hydraulic tomography. *Advances in Water Resources*, 82, 83-97.

Sun A. (2018). Discovering State-Parameter Mappings in Subsurface Models Using Generative Adversarial Networks. *Geophysical Research Letters*, 45(20), 11137-11146.

Sun N.-Z. & Yeh W.-G. (1990). Coupled inverse problems in groundwater modeling: 1. Sensitivity analysis and parameter identification. *Water Resources Research*, 26(10), 2507-2525.

Tahmasebi P., Kamrava S., Bai T. & Sahimi M., 2020. Machine learning in geo- and environmental sciences: From small to large scale. *Advances in Water Resources*, 142, 103619.

Tarantola A. & Valette B. (1982). Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics*, 20(2), 219-232.

Vu M.T. & Jardani A. (2021). Convolutional neural networks with SegNet architecture applied to three-dimensional tomography of subsurface electrical resistivity: CNN-3D-ERT. *Geophysical Journal International*, 225(2), 1319-1331.

Vu M.T., Jardani A., Krimissa M., Fischer P. & Ahfir N. (2019). Hydraulic tomography in time-lapse mode for tracking the clogging effects associated with the colloid injection. *Advances in Water Resources*, 133, 103424.

Wei Z. & Chen X. (2019). Deep-Learning Schemes for Full-Wave Nonlinear Inverse Scattering Problems. *IEEE Transactions on Geoscience and Remote Sensing*, 57(4), 1849-1860.

Wu Y. & Lin Y. (2018). InversionNet: A Real-Time and Accurate Full Waveform Inversion with CNNs and continuous CRFs. *arXiv*:1811.07875.

Xu T. & Gómez-Hernández J. (2018). Simultaneous identification of a contaminant source and hydraulic conductivity via the restart normal-score ensemble Kalman filter. *Advances in Water Resources*, 112, 106-123.

Yeh J. T.-C. & Liu S. (2000). Hydraulic tomography: Development of a new aquifer test method. *Water Resources Research*, 36, 2095-2105.

Yeh T.-C. J. & Zhu J. (2007). Hydraulic/partitioning tracer tomography for characterization of dense nonaqueous phase liquid source zones. *Water Resources Research*, 43(6).

Yoon H., Hyun Y. & Lee K.-K. (2007). Forecasting solute breakthrough curves through the unsaturated zone using artificial neural networks. *Journal of Hydrology*, 335(1), 68-77.

Zarita Z. & Ong P. (2008). Function Approximation Using Artificial Neural Networks. *WSEAS Trans. Math.*, 7(6), 333–338.

Zha Y., Yeh T.-C. J., Illman W. A., Onoe H., Mok C. W., Wen J.-C., Huang S.-Y. & Wang W. (2017). Incorporating geologic information into hydraulic tomography: A general framework based on geostatistical approach. *Water Resources Research*, 53(4), 2850-2876.

Zhang G., Eddy Patuwo B., & Hu Y. M. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14(1), 35-62.

Zhang Z. & Lin Y. (2020). Data-Driven Seismic Waveform Inversion: A Study on the Robustness and Generalization. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10), 6900-6913.

Zhu J., Cai X. & Jim Yeh T.-C. (2009). Analysis of tracer tomography using temporal moments of tracer breakthrough curves. *Advances in Water Resources*, 32(3), 391-400.

Zio E. (1997). Approaching the inverse problem of parameter estimation in groundwater models by means of artificial neural networks. *Progress in Nuclear Energy*, 31(3), 303-315.