



HAL
open science

Mapping discrete fracture networks using inversion of hydraulic tomography data with convolutional neural network: SegNet-Fracture

M. T. Vu, Abderrahim Jardani

► **To cite this version:**

M. T. Vu, Abderrahim Jardani. Mapping discrete fracture networks using inversion of hydraulic tomography data with convolutional neural network: SegNet-Fracture. *Journal of Hydrology*, 2022, 609, 10.1016/j.jhydrol.2022.127752 . insu-03661809

HAL Id: insu-03661809

<https://insu.hal.science/insu-03661809>

Submitted on 22 Jul 2024

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

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



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Mapping discrete fracture networks using inversion of hydraulic**
2 **tomography data with Convolutional Neural Network: SegNet-Fracture**

3 M. T. Vu ¹, A. Jardani ¹

4 ¹ Université de Rouen, M2C, UMR 6143, CNRS, Morphodynamique Continentale et Côtière,
5 GeoDeepLearning consortium, Mont Saint Aignan, France

6 **Abstract:** In this paper, we propose a new method to map the fracture network structure in a
7 heterogeneous aquifer from inversion hydraulic head data measured during pumping tests in hydraulic
8 tomography mode. This inversion tool is based on the new concept of convolutional neural networks,
9 which provides a direct approximation to the inverse function linking fracture geometry to hydraulic
10 data. In order to handle the highly nonlinear inverse function more effectively, an advanced neural
11 network is developed from Segnet architecture with encoder-decoder structure, which excels in image
12 processing to translate the water level image associated with the pumping tests at the input into a
13 fracture map at the output. The network is trained with a synthetic dataset where the fracture structure
14 and matrix heterogeneity are randomly generated, and the hydraulic head are obtained by solving the
15 groundwater flow equation. The trained network accurately maps different complexity levels of
16 fractures embedded in a matrix with heterogeneous transmissivity.

17 As a data-driven approach, the accuracy of the mapping depends on the quantity, quality, and
18 relevance of the synthetic dataset used in the training phase. While generating data to train the
19 network requires effort, the trained network performs each inversion instantly. The inversion result
20 appears to be stable even in the presence of data noise, reliably interprets the hydraulic data if they
21 share comparable fracture and matrix properties as specified in the training models.

22

23 *Keywords: Neural Network, Fractured aquifer, Convolution Neural Network, Groundwater,*
24 *Inversion.*

25

Introduction

27 Characterizing the hydraulic properties of fractured/karstic aquifers is known to be one of the most
28 challenging tasks in the quest to understand groundwater dynamics and contaminant transport in these
29 complex environments (Neuman, 2005). Indeed, due to the large variability in fractures and rock
30 matrix, as well as the permeability gap between them, these fractured aquifers exhibit extremely
31 complicated groundwater flow patterns. Water flows, in this heterogenous structure, are mostly
32 concentrated in a percentage of the highly permeable matrix and in sparse fracture networks that
33 emerge across a vast impermeable region (Budd & Vacher, 2004; Maréchal et al., 2004).
34 Identification of these fracture networks is therefore crucial in determining strategies for exploiting
35 and protecting any reservoir.

36 Various tools have been tested and developed in the literature to locate preferential flows associated
37 with the presence of fractures, including tracer tests, conventional pumping and cross-pumping
38 techniques (Dverstorp et al., 1992; Dahan et al., 1999, Dausse et al., 2019). Cross-pumping tests are
39 referred to as a hydraulic tomography technique that captures the spatial variability of the hydraulic
40 properties of the porous or fractured aquifer from joining hydraulic drawdown data recorded in
41 successive pumping tests (Illman, 2014). Typically, processing of these acquired hydraulic data is
42 performed using an inversion code to determine models of hydraulic conductivity and storativity
43 which are numerically consistent for the observed hydraulic data. The inversion method employs a
44 forward problem to numerically provide hydraulic responses for a given model of hydraulic properties
45 by solving for groundwater flow based on the Darcy equations. In this method, the forward operator is
46 solved repeatedly until achieving a satisfactory match between the observed and simulated hydraulic
47 data (Yeh, 1986).

48 For the characterization of fractured/karstic aquifers, the groundwater flow equation can be solved
49 using one of two distinct approaches: Equivalent Porous Medium (EPM) and discrete fracture
50 networks (DFN). In the ECM concept, fracture networks are approximated as a porous medium in
51 which the fractured aquifer is fully discretized with small cells that represent the effective values of
52 transmissivity and storativity that are being identified during the inverse process (Yeh & Liu, 2000;

53 Zha et al., 2015; Tiedeman & Barrash, 2020). This approach is straightforward to implement and
54 parameterize the spatial heterogeneities of the hydraulic property fields in the inverse process. It,
55 however, ignores the discontinuity features of groundwater flow in fractured environments (Wang et
56 al., 2016). In contrast, DFN concepts preserve the fracture shape and its hydraulic discontinuities via
57 parameterizing the fracture and conduit structures as lines in 2D or surfaces in 3D and linking their
58 hydraulic properties to their apertures (Fischer et al., 2018; Ringel et al., 2019; Mohammadi & Illman,
59 2019). Despite its realistic representation, this parameterization is practically complicated due to the
60 significant time required in numerical simulation and the difficulties in determining both the spatial
61 geometries of the fractures and their apertures in the inverse process. For this reason, the number of
62 applications of DFN parameterization in the inverse problem with stochastic or deterministic
63 optimizers remains limited in comparison to EPM.

64 In this paper, we explore the potential of deep learning tools in the inversion of hydraulic data
65 monitored in pumping tests to identify conduit and fracture networks. Compared to stochastic and
66 deterministic algorithms, the deep learning algorithm performs a direct approximation of the inverse
67 function: linking the piezometric measurements as input to the hydraulic properties in the output. This
68 link between the two is made through multiple linear and nonlinear operations performed sequentially
69 on layers of the network. Such a decomposition provides an efficient way to approximate highly
70 nonlinear functions as in the case of inversion operators (Elanayar & Shin, 1994). The concept
71 involves the use of several multidimensional coefficients, called weights and biases, assigned to the
72 network layers. The determination of these coefficients is based on an optimization process performed
73 during the training operation, which uses a known set of input and output data, called the training
74 dataset. Thus, most of the effort and computational time required to build a neural network relates to
75 building the database for training and the training task itself. Once trained, however, the trained
76 network provides an end-to-end operator that performs inversions of experimental data instantly.

77 Deep learning algorithms are progressively gaining ground in the geosciences, particularly the
78 emergence of the new powerful neural networks that enables to handle complex problems in

79 geophysical imaging, remote sensing and other environmental monitoring (Benjamin et al., 2020;
80 Lary et al., 2016; Moseley & Krischer, 2020). Some preliminary investigations adopted deep learning
81 algorithms using vanilla neural networks, for example, to interpret measurements from pumping or
82 tracing tests (Aziz & Wong, 1992; Zio, 1997; Balkhair, 2002; Trichakis et al., 2009). Comparative
83 analyzes reveals the efficiency of neural network as a replacement for traditional interpretations.
84 However, fully connected layers in a vanilla neural network with a high number of learnable
85 coefficients lead to training task overload when dealing with high-dimensional problems. Recent
86 breakthroughs in the field of data science have given rise to the new networks based on the
87 convolutional concept that efficiently tackle tomography problems with reasonable effort (LeCun &
88 Bengio, 1998). Convolutional neural networks (CNN) translate the input image by scanning the
89 shared small filters region-by-region to extract features on the entire image and construct local
90 connections between neighboring pixels. Such a sharing mechanism facilitates the training task and
91 information extraction compared to processing point-by-point as in the vanilla network (Rawat &
92 Wang, 2017).

93 Convolutional architecture performance has sparked a surge of research and applications in
94 geophysical and hydraulic imagery for approximating inversion operators to identify the subsurface
95 properties from sparse observations on the boreholes or at the surface. Among them, Liu et al. (2020)
96 built a CNN architecture to map the 2D electrical resistivity in the subsurface from apparent resistivity
97 images inferred from surface data, while Vu & Jardani (2021) developed a CNN-ERT3D neural
98 network to process 3D resistivity tomography. Other efforts applied the networks in seismic
99 exploration to reconstruct the seismic velocity models directly from seismogram data (Li, et al., 2020;
100 Zhang & Lin, 2020, Apolinario et al., 2019; Park & Sacchi, 2020). Similar ideas were tested for the
101 inversion of electromagnetic data to identify the 1D or 2D subsurface resistivity (Puzyrev &
102 Swidinsky, 202; Puzyrev, 2019)

103 In hydrogeology, Laloy et al accessed the effectiveness of a generative adversarial network (GAN)
104 in mapping hydraulic conductivity heterogeneities in 2D and 3D binary pattern (Laloy et al., 2017;
105 Laloy et al., 2018). Bao et al. (2020) extended the coupling of GAN and Ensemble Smoother with
106 Multiple Data Assimilation, to reconstruct the binary media and reduce the prediction uncertainty.

107 Sun (2018) suggested a composition of GAN and state parameter identification to map simultaneously
108 the parameter space, as the permeability, and the corresponding model state, as the subsurface
109 structure. Other initiatives employed encoder-decoder architectures to image the spatial distribution of
110 the transmissivity field from measurements of hydraulic heads (Jardani et al., 2022) or tracer
111 concentrations (Vu & Jardani, 2022). However, all these applications focus solely on mapping
112 heterogeneity in a porous aquifer; to our knowledge, only a few works contribute to determination of
113 the fracture network geometry.

114 In this paper, we propose a CNN based on the SegNet architecture to identify the fracture network
115 structure from synthetic data of hydraulic head monitoring in a fractured aquifer. The study is
116 structured as follows: Section 2 outlines the theoretical background which summarizes the forward
117 problem with groundwater equations, the concept of the inverse problem approximated by Segnet and
118 its differences from classical inversion methods. Details of the data generation, learning process and
119 validation of the trained network follows in Section 3. Section 4 is devoted to accessing the
120 performance of the proposed network in the face of possible interferences in practice, which could be
121 related to the data quality and quantity, the heterogeneity in matrix, or the relevance of dataset
122 features. Finally, a summary of the main findings with suggestions for future work concludes the
123 paper.

124 **2. Theoretical Background**

125 In this section we summarize the theory of which the first part includes the forward problem, the
126 concept of the inverse problem with neural network, followed by the description of proposed neural
127 network architecture and its working mechanisms.

128 **2.1 Forward problem: Groundwater modelling**

129 As mentioned earlier, inversion by a deep learning algorithm is also based on the forward problem
130 where generating the training data set involves solving the groundwater flow equation. This training
131 data consists of the numerically computed hydraulic response to the pumping tests corresponding to a
132 given synthetic transmissivity field and discrete fracture networks. Fractures in the aquifers are
133 geometrically represented by 1D lines and assigned to an equivalent transmissivity, whereas matrix
134 rock is characterized as a porous medium - a heterogeneous transmissivity field. The neural network
135 then learns the relationship between the hydraulic head responses and the permeable fracture
136 structures rooted in the training dataset. In this work, the inversion task aims only to reconstruct the
137 geometric structure of fractures without rebuilding the transmissivities in fracture and matrix.

138 In steady-state conditions, numerical simulation of groundwater flow in a fractured aquifer is
139 performed using two connected equations derived from Darcy's Law that describe water flows in the
140 rock matrix and fractures (Kohl et al., 1997):

$$141 \begin{cases} \nabla(-T_m \nabla h) = q & \text{in the matrix } \Gamma_m \\ \nabla_T(-T_f \nabla_T h) = q & \text{in the fracture } \Gamma_f, \end{cases} \quad (1a)$$

142 subjected to the following boundary conditions

$$143 h = h_D \text{ at } \Gamma_D, \quad (1b)$$

144 where ∇ and ∇_T are the gradient and tangent gradient operators, respectively; T_m and T_f [m²/s]
145 denote the transmissivity of heterogenous matrix (Γ_m) and fractures (Γ_f), respectively. The fractures
146 are shaped as lines which is randomly generated, while their equivalent transmissivity is assumed
147 constant. h [m] is the hydraulic head; q [l/s/m] represents the punctual water source at the
148 extraction/injection wells. To form boundary conditions, the observation area is embedded in a large
149 buffer zone assigned to the mean transmissivity of the matrix blocks. Such large bounds limit the
150 boundary effects on the computational results. For completeness, a constant hydraulic head (h_D) is
151 imposed at the buffer zone periphery (Γ_D).

152 In this study, the source of water is the paired pumping wells, each pair consists of one extraction
153 and one injection well located on opposite sides of the observation area. Each paired well pump at the

154 same constant rate over the operating period. Finally, we solve the set of flow equations numerically
155 using a finite element method in COMSOL software to determine the hydraulic head at observed
156 points, which are then taken as input to the neural network.

157 **2.2 Inversion problem with Neural network: Mapping fracture from groundwater observation**

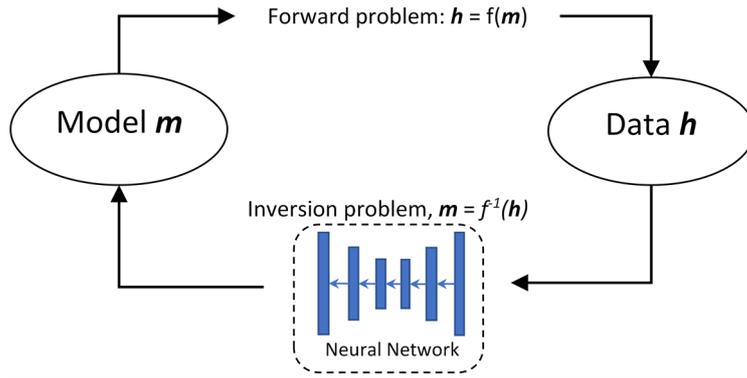
158 As mentioned in the previous problem, equation (1) represents a numerical operator for computing
159 piezometric data in response to a pumping test in an aquifer with matrix transmissivity T_m , fracture
160 transmissivity T_f with geometry C_f . The operator can be expressed as the following function with three
161 dependent parameters:

$$162 \quad \mathbf{h} = f(T_m, T_f, C_f), \quad (2)$$

163 The operator works as a data generator for training the neural network which performs an inverse
164 function to determine the fracture geometry from the hydraulic head data (see Figure 1). Here we
165 focus only on reconstruction of the fracture geometry, without identifying the transmissivity field of
166 matrix rock and fractures which will be handled in a future study. Thus, the aim of this study is to
167 retrieve the fracture geometry from the hydraulic data by approximating the nonlinear inverse
168 operator f^{-1} between two fields using Segnet network based on thousands of synthetic data sets:

$$169 \quad C_f = f^{-1}(\mathbf{h}) = \Psi(\mathbf{h}, \Theta), \quad (3)$$

170 where Ψ represents the neural network, whose learnable parameters are denoted as Θ .



171

172

173

174

175

Figure 1: A scheme representing the prediction of fracture geometries from the hydraulic head data using Neural network. The approach consists of the learning phase: solving the forward problem to generate the training data, and the use of neural network to establish the relationship between the fracture geometry and the hydraulic data from the training models.

176

177

178

179

Indeed, the neural network, such as Segnet, offers a possibility to directly relate the hydraulic data to the fracture geometries by determining a set of learnable parameters defined as sequential linear and nonlinear operations in Segnet layers. Determination of the parameters is performed from solving an optimization problem with the following objective function

180

$$\Theta = \operatorname{argmin} \left\{ \sum_1^N \|C_f \sim \Psi(h_i, \Theta)\| \right\}, \quad (4)$$

181

182

183

where the pair (C_f, h_i) denotes a model and its corresponding measurement, N is the number of pairs in the training dataset; the operator \sim denotes the cross entropy to measure the difference between the truth and predicted models.

184

185

186

A training process often demands a large dataset, of which the generation task takes the most time in the building of a neural network. However, once the training is completed, the trained network performs an inference instantly, an end-to-end operator without intervention from users.

187

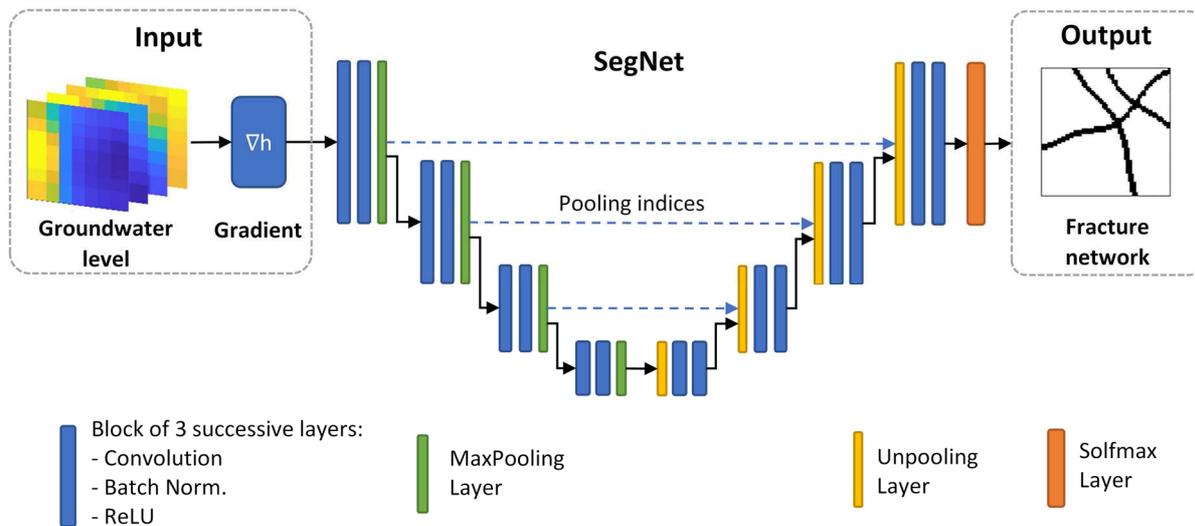
2.3 SegNet Architecture - Direct inversion function

188

189

In this paper, the inversion operator is approximated by the SegNet architecture which is designed to efficiently locate the fracture map from groundwater level measurements relied on the concept of

190 convolutional neural network (see Figure 2). The SegNet architecture was developed by
 191 Badrinarayanan et al. (2017) and has been rapidly proven in a variety of domains with numerous
 192 applications, including: traffic scene detection (Kendall et al., 2017; Jiang et al., 2020; Qin et al.,
 193 2020), satellite image processing (Khryashchev et al., 2018; Mohammed & Edward, 2019; Sariturk et
 194 al., 2020), geophysics (Pham et al., 2018; Mukhopadhyay & Mallick, 2019; Vu & Jardani, 2021).
 195 However, it is still uncommon in hydrology, as well as other advanced algorithms in this discipline,
 196 where progress is modest compared to other geosciences such as geophysics.



197

198 **Figure 2: SegNet architecture consists of 4 encoder-decoder levels, with 63 neural layers to**
 199 **interpret the fracture network geometry from groundwater level measurement.**

200 Typically, a neural network consists of multiple layers of artificial neurons that mimic the
 201 operations of their biological counterpart through mathematical functions (Indiveri et al., 2011). When
 202 an input image is introduced into the network, each layer operates on activation functions to interpret
 203 features before passing them on to the next layer. The process ends with a mapping result
 204 reconstructed from the learned features. By matching the truth in the given datasets with its
 205 reconstruction from the network, all parameters in the activation functions are determined, which is
 206 then named as the training task. Network performance evidently relies on how efficiently the features

207 are translated after this training process. A more complex network appears to better refine the learned
208 features that is why the proposed neural network is based on an advanced architecture (Bengio et al.,
209 2013).

210 The proposed architecture, as shown in Fig.2, consists of successive network blocks with four
211 encoders and four decoders, finalized by a pixel-wise classification layer to map the fractures. This U-
212 shaped form consists of 63 layers divided into two paths: a contracting one with encoders and an
213 expanding one with decoders, each having the same structure but arranged in reverse order. The
214 encoders extract the features embedded in the input images using a set of sequential convolutions,
215 each followed by a batch-normalization (BN) and a rectified linear unit (ReLU), and finally a max-
216 pooling operation. On the opposite path, the decoders start with an un-max-pooling operation
217 followed by a sequence of convolutions, BN and ReLU layers. The max-pooling indices from each
218 encoder are assigned to the corresponding decoder to compensate for the resolution loss in the
219 encoders. This sharing mechanism is the core aspect of a SegNet architecture (Badrinarayanan et al.,
220 2017).

221 But how does the network reconstruct the fractures? The concept of a convolutional neural network
222 relies on scanning small filters over entire images to extract key features which represent the local
223 connectivity among neighboring pixels. In this sequential convolution, the first layer usually extracts
224 basic features, such as horizontal or diagonal edges, while more complex features are detected in the
225 next layers, such as corners, edges, or shapes (Ghafoorian et al., 2017). The deeper it goes, the more
226 complex features the network can detect (Fukushima, 1988). In this manner, convolutions translate
227 the fracture information embedded in the input water level images to yield the fracture map.

228 Number of convolution level is designed in response to problem complexity; however, deeper
229 convolutional translation inevitably entails a large system of filters that soon overburdens the training
230 task. To save time and resources, a max-pooling operation is invoked to sort out only maximum
231 values. By shrinking the spatial size of the convolved features, the operation speeds up the processing
232 and is effectively deals with overfitting issue. By working only on maxima, the max-pooling
233 conserves the most prominent invariant features along with dimensionality reduction (Nagi et al.,
234 2011). Since the fracture is prominent in this case, the operation acts as a denoising agent to highlight

235 fractures from the matrix heterogeneity. However, the operation results in a downsized output and
236 lose of information. To compensate the side effects, state of each max-pooling is transferred to the
237 corresponding un-max-pooling in decoders to counterbalance the degradation.

238 Thanks to sharing max-pooling indices, the SegNet outperforms other networks such as FCN,
239 DeepLab, DeconvNet in terms of accuracy (Badrinarayanan et al., 2017). For example, the proposed
240 network with 4 encoder-decoder levels requires only 0.5 million parameters, which is far less than
241 Unet with 8 million. A smaller size is also related to the fact that the network employs only $3 \times 3 \times 64$
242 filters which are partly smaller than those in the Unet. Recall that our network is designed with 4
243 encoder-decoder levels, shallower than the original network as it processes fewer output segmentation
244 labels.

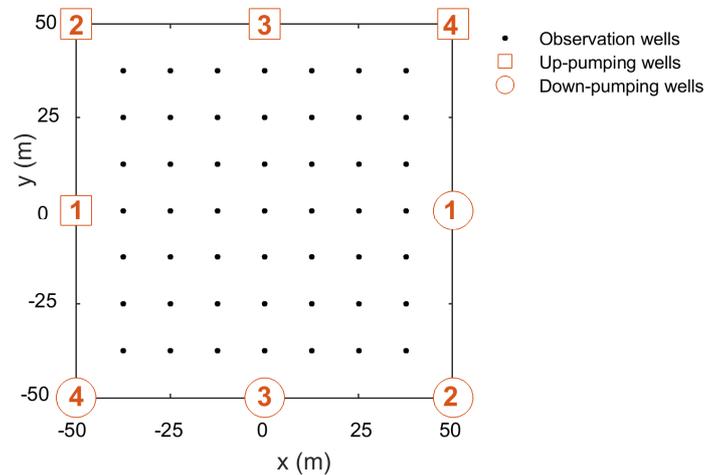
245 **3. Application**

246 This section is devoted to the training process in a synthetic application. We begin with a
247 generation of aquifer models, each with a fracture network embedded in a heterogenous ground.
248 Water level responses in each aquifer model are observed with a monitoring scheme designed to
249 mimic a real experiment, where the acquisition setup maps the groundwater level in injection tests.
250 Following that, the acquired data is processed to feed training of the network. The training is then
251 implemented, along with hyperparameters, evaluation metrics, and its evolution during the process.
252 Typical results are shown to validate the trained network.

253 ***3.1 Aquifer model: fracture network embedded in a heterogeneous ground***

254 In this study, the heterogeneity in aquifer transmissivity is generated based on statistical parameters
255 following a Gaussian distribution. For this purpose, we use the SGeMS code implemented in Matlab
256 to create 35,000 transmissivity models, where the distribution of $\log_{10}T$ is randomly constructed using

257 a Gaussian variogram with constant mean and variable range (Remy et al., 2009). Heterogeneity in
 258 transmissivity models ranges from 10^{-8} to 10^{-4} m²/s in four orders of magnitude which corresponds to
 259 the range of a permeable aquifer. The generated transmissivity fields are then assigned to a 100×100
 260 m area with a locally constant aquifer depth (Figure 3). Each ground model is assigned a fracture
 261 network geometry which consists of no to three fractures assigned randomly. The transmissivity of
 262 fracture is set at 10^{-2} m²/s which is of orders conductive than the surrounding matrix with a mean of
 263 10^{-6} m²/s. All fractures are formed from 1D random generation using a Gaussian variogram in
 264 SGeMS, and then randomly rotated to redirect their orientation.



265

266 **Figure 3: Setting of monitoring area which consists of 49 observation wells arranged in a**
 267 **regular array (black dots) and 4 paired pump wells on the boundaries implemented successively**
 268 **(red numbers).**

269 To establish a monitoring scheme, we install 49 observation wells in a regular array in both
 270 directions on the synthetic aquifers (Figure 3). At the field boundaries, we arrange 4 paired pump
 271 wells, in each pair an up-pumping well is in opposite position to its down-pumping. This arrangement
 272 of the hydraulic forces in pairs allows realizing a hydraulic disturbance of the whole investigation
 273 area. In addition, this configuration in practice saves water resources during the hydraulic tomography
 274 test.

275 **3.2 Data acquisition and data processing**

276 In each model, groundwater responses due to pumping/injection tests are determined by computing
277 the forward problem described in Section 3.1. Four pumping/injection tests are performed
278 sequentially at a constant flow rate of 10 L/min in each well. Hydraulic measurements are acquired on
279 the 49 observation wells and then interpolated using the Akima method to obtain 4 maps of hydraulic
280 head.

281 The input and output maps in the training datasets use the same spatial resolution, being discretized
282 by a 64×64 grid on the 100×100 m study area. Recall that in this study, we focus only on mapping
283 the network of fractures as preferential flow paths during pumping/injection tests. The presence of
284 fractures is often associated with abrupt changes in groundwater flow. To better identify these
285 discontinuities, we calculate the gradient in both directions (x and y) of the hydraulic head maps
286 interpolated from the hydraulic head measurements. The resulting gradients are then reshaped into a
287 64×64 grid. Each pumping/injection test generates 2 gradient maps (x , y) to build the gradient map
288 set of 8 channels ($64 \times 64 \times 8$) as the input of neural network training. For the output map, the fractures
289 are labelled as 1 and the others as 0 on the gridded map.

290 To quantitatively assess the predictions accuracy of the network, we use the accuracy coefficient as
291 defined below

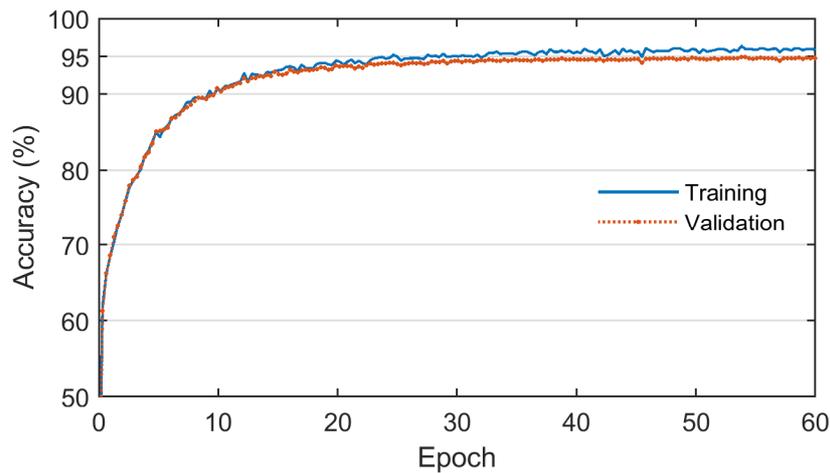
$$292 \quad Accuracy = \frac{1}{N} \sum_{j=1}^N (y_j == \hat{y}_j), \quad (5)$$

293 where y and \hat{y} denote the labels (fracture or ground) in the true model and its corresponding
294 prediction, respectively. N represents the total pixel number in the output image (64×64).

295 **3.3 Training process and results**

296 For the training phase, we randomly separate 25,000 aquifer models with corresponding measured
297 data: 20,000 for training the network and 5,000 for validation during the process. A high fraction of

298 validation dataset aims to better evaluate efficiency of the training process. Training is performed
299 using the ADAM optimization algorithm implemented in MATLAB, on a Dell Precision Tower 5810
300 with a single GPU NVIDIA Quadro K2200. The task is completed in 5 hours, 60 epochs with a
301 constant learning rate of 0.01 and a batch size of 128. Figure 4 summarizes evolution of accuracy
302 along the training process.

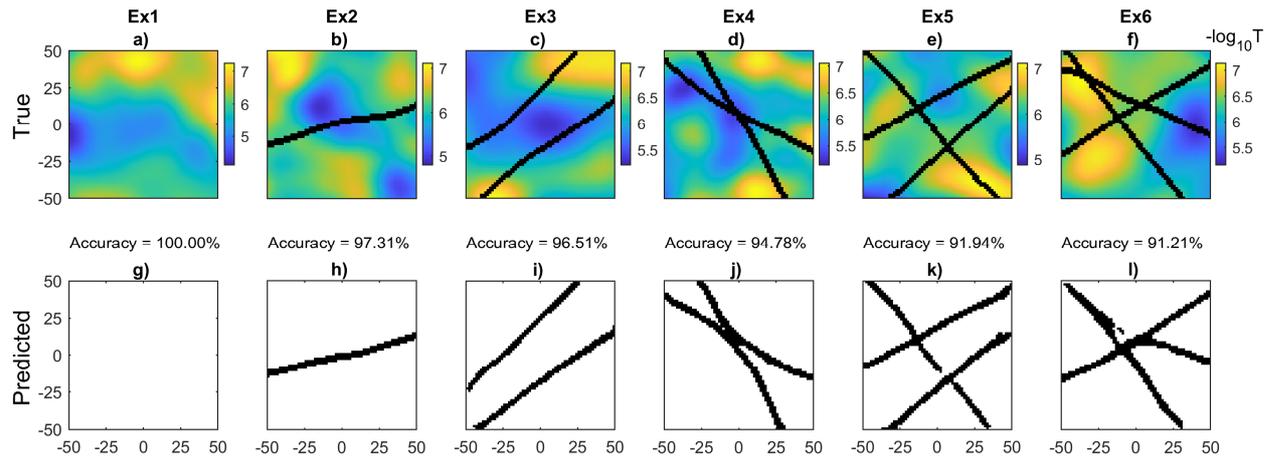


303

304 **Figure 4: Accuracy evolution in the training process, including 20,000 data sets for training**
305 **and 5,000 data sets for validation.**

306 The trained network is then evaluated with 10,000 unseen models, each model is executed in 0.006
307 seconds, almost instantly. Fractures are efficiently mapped over 10,000 samples with an average
308 accuracy of 94.83%. The overall reconstruction quality is summarized in the histogram shown in
309 Figure 11. High accuracy prediction is related to the fact that fractures with higher conductivity drive
310 dynamics in the aquifer that favor their detection. To investigate the sensitivity of the proposed model,
311 we train the network 50 times with the same dataset and hyperparameters as above. The results show
312 a consistent solution, where the average accuracy of the testing dataset centers at 94.86 %, with a
313 standard deviation of 0.1 % over the trained models.

314 To discuss the accuracy of the reconstructions in detail, we select and show six representative
315 results of the tested models in Figure 5. Six showcases are arranged in different levels of complexity
316 of the fracture distribution, starting with the models without fractures and ending with the networks
317 consisting of three fractures.



318

319 **Figure 5: Six representative models are selected to illustrate the effectiveness of the trained**
 320 **network (Case 1). From top to bottom by row: the real, accuracy, and predictions; from left to**
 321 **right by column: the representative cases from Ex1 to Ex6. Fractures are efficiently mapped**
 322 **with a high accuracy. The accuracy of reconstruction relies on the density and complexity of the**
 323 **fracture network overall.**

324 The accuracy of reconstructions overall refers to a clear dependency on the density of the mapped
 325 fractures and their complexity level. Without fractures, Ex1 shows an accurate prediction where the
 326 matrix transmissivity heterogeneities do not produce artifacts in the interpretation. A further
 327 discussion on this ground impact is addressed in Section 4.5 with higher heterogeneity. In the
 328 following cases, Ex2 and Ex3, the trained network also properly reproduces the models with one or
 329 two fractures. However, in a more complex case in Ex4 where two fractures cross closely, there is a
 330 minor misinterpretation with a fused joint. Even though, the reconstruction still faithfully represents
 331 the main aspects of the fractures.

332 Case Ex5 otherwise presents a dense fracture network in which three fractures are oriented in
 333 different directions; these patterns are somewhat more complex than the previous ones, but the
 334 network is still able to identify the fracture paths with accuracy. Along the reconstructed fractures in
 335 Figure 5hij, other slightly deformed curves or a minor discontinuity can be observed, possibly due to

336 the influence of ground heterogeneity. The ground heterogeneity exhibits less influence on subsurface
337 flow compared to fractures. However, by introducing more noise and locally changing the water table,
338 heterogeneity may misguide the understanding of fracture networks. A detailed discussion on this
339 problem is further elaborated in Section 4.1 for results from homogeneous models. In the last case,
340 Ex6 also consists of three fractures, but with a very complex geometry, including two fractures close
341 to each other, which the network cannot distinguish in the reconstruction. This misinterpretation stems
342 from the fact that the two fractures behave as similar as a single fracture in the middle. In general,
343 however, the fracture reconstructions are very satisfactory.

344 Indeed, number of interferences can impact prediction efficiency which may occur in a real field
345 under diverse situations, some of the key issues are addressed in the following sections.

346 **4. Discussion**

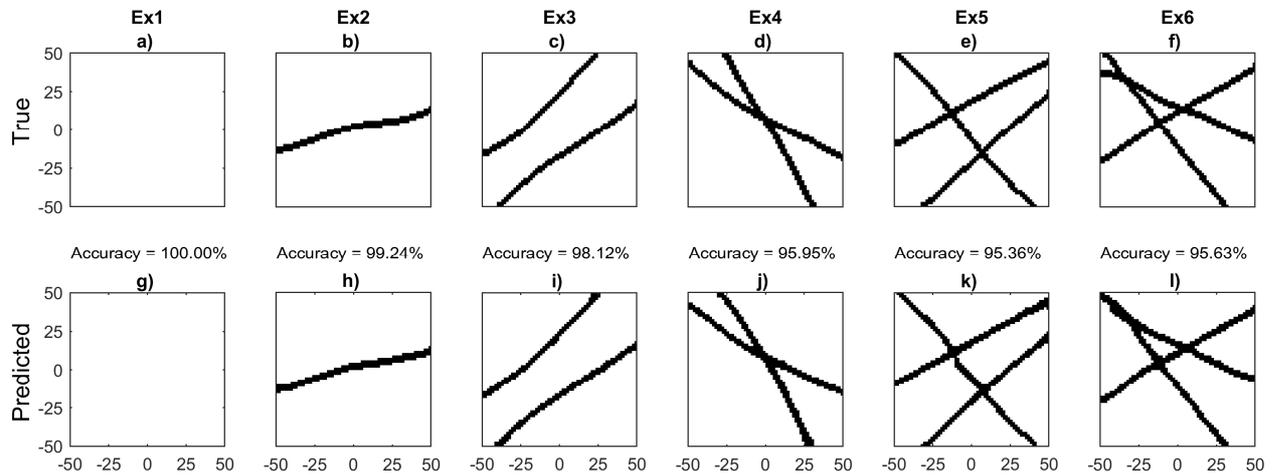
347 This section is devoted to studying the impact of various sources of uncertainty that can affect the
348 quality of the reconstructions with the network, such as: the uncertainty about the degree of
349 heterogeneity of the ground transmissivity and the choice of the size and the nature of features of the
350 training models. We also study the impact of the number of piezometers used in hydraulic
351 tomography and the noise that can alter these hydraulic measurements.

352 ***4.1 Effect of heterogeneity in the matrix***

353 Both porous ground and fracture network geometry associates in forming the subsurface dynamics;
354 yet the roles of each system in driving the groundwater are distinct. Dominant conductivity and
355 connectivity of fracture network make it superior to the surrounding porous ground. While fractures
356 regulate the flow regime, the permeable ground, even of a lesser order, is equally important.

357 To investigate the influence of heterogeneity in matrix transmissivity, we repeat the learning
358 process described in Section 3, but this time with homogeneous transmissivity models. We regenerate
359 35,000 models using the same fracture network configuration embedded in ground with a constant

360 transmissivity of 10^{-6} m²/s. The network is re-trained using the same piezometric configuration and
 361 network hyperparameters as in Section 3.3. The overall average accuracy for 10,000 test models
 362 shows a significant improvement from 94.83% in Case 1 to 96.94% in this case. Details of the
 363 representative models are shown in Figure 6, and the histogram is summarized in Figure 11.



364

365 **Figure 6: Examples of reconstruction with a homogeneous ground transmissivity of 10^{-6} m²/s**
 366 **(Case 2). From top to bottom by row: the truth, accuracy, and predictions; from left to right by**
 367 **column: the representative cases from Ex1 to Ex6. Without ground heterogeneity, fractures are**
 368 **reconstructed with a better accuracy. Complexity in the fracture network is better delineated**
 369 **than in the previous case which prove the impact of the ground heterogeneity on overall**
 370 **accuracy.**

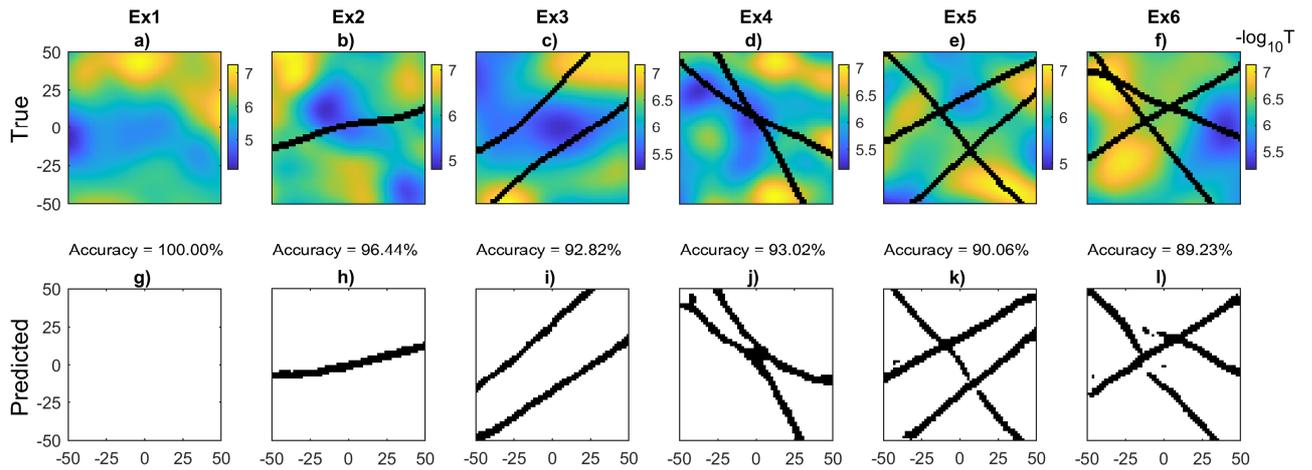
371 The showcases in Figure 6 prove a better reconstruction due to the absence of disturbance effect
 372 associated with a heterogenous ground, particularly for the complex configurations in Ex4 and Ex6.
 373 The fused bottleneck in the prediction of Ex4 has clearly disappeared, accompanied by a better shape
 374 for both fractures. Other slightly deformed curves that appeared in Figure 5hij also do not appear in
 375 Figure 6hij. Similarly, all the fractures are well shaped overall in Ex6 where the two closed fractures
 376 are clearly delineated, even some minor disturbances are shown on the top of the fractures.
 377 Comparison of the accuracy between the two sets of reconstructions confirms the heterogeneity

378 importance of the porous ground in determining the fracture network structure. Further discussion on
379 this issue is developed in Section 4.5 where the heterogeneity deviates from a Gaussian distribution
380 predefined in the training dataset.

381 *4.2 Effect of dataset size*

382 In general, the effectiveness of a trained neural network in a deep learning algorithm depends
383 strongly on the size of the dataset used in the learning phase. In our inverse problem, generating the
384 training dataset involves solving a mass of a forward problem with numerical tools, which is the
385 longest and most laborious phase in building a neural network. The ideal strategy is to determine the
386 optimal size which reduces computation effort while resulting in an accurate prediction. This can be
387 accomplished by gradually increasing the data size and periodically verifying the prediction accuracy
388 in the training, validation and testing phases until a satisfactory result is achieved.

389 In this section we examine the impacts of training data size on driving the reliability of
390 reconstructions by analyzing the predictions obtained with networks trained with three different sizes:
391 5000, 10000, 20000 and 30000. The networks formed with these datasets are tested on the 10000
392 unseen models as in Section 3.3, of which the metric results are reported in Table 1 and in Figure 11
393 for the histogram. We recall the six representative models as in the previous tests to analyze the
394 impact of dataset size in detail (see Figure 7).



395

396 **Figure 7: Results from the network trained from 5,000 models in the dataset (Case 3). From**
 397 **top to bottom by row: the truth, accuracy, and predictions; from left to right by column: the**
 398 **representative cases from Ex1 to Ex6. The accuracy of fracture reconstruction relies on the size**
 399 **of training dataset. A small training dataset results in a degradation in the prediction accuracy**
 400 **with artifacts or missing feature in the reconstruction. Larger dataset enhances the prediction**
 401 **quality, an asymptotic enhancement.**

402 The use of a minimal number of training data is sufficient to results in reliable maps when the
 403 reconstructed fracture networks are straightforward, as in the case of models Ex1 to Ex3. However,
 404 when the configurations are difficult, as in the cases of models Ex4 to Ex6, the predictions do not
 405 provide accurate identifications of the fracture structures with minor artifacts and missing features in
 406 the reconstruction. A larger training dataset assists improving reconstructions in enriching the learning
 407 process through introducing more complex features (Advani et al., 2020). Better generalization thus
 408 requires a wide range of alternative models to cover the high complexity in karstic aquifer and to
 409 tackle the overfitting issue.

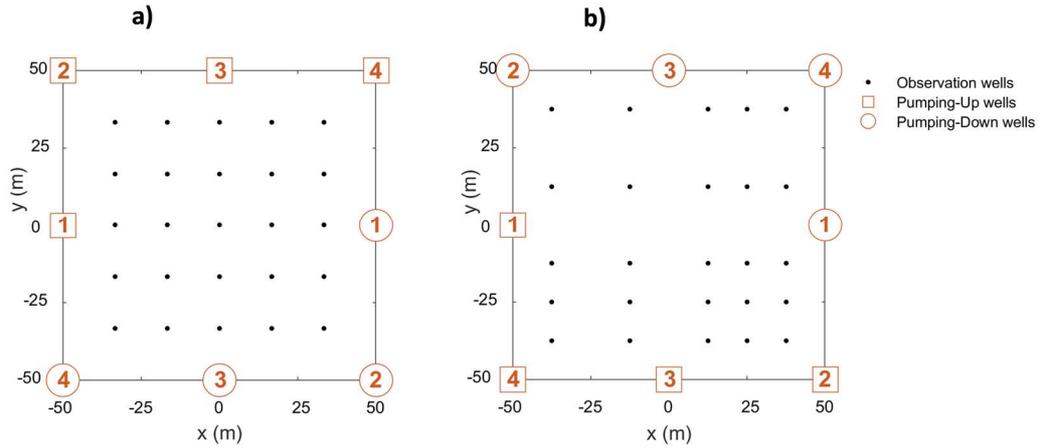
410 Details of prediction accuracy in Table 1 reveal a clear dependence on the size of dataset, with
 411 overall mean accuracy increasing from 93.08% for the first subset (5,000 models in training), to
 412 93.95% for the second (10,000 models in training), and to 95.19% of the third (30,000 models in

413 training). The enhancement in accuracy is, however, not proportional to the enlargement of dataset
414 size since results from subsets of 20,000 (Case 1) and 30,000 models show no clear difference (with
415 0.3% of improvement in accuracy for extra 10,000 samples). Overfeeding data overburdens the
416 training task since increasing data volume ineffectively contribute to improve accuracy when
417 performance reaches an asymptotic stage (Amari et al., 1997). As mentioned at the beginning of this
418 section, the choice of training data size depends on the desired accuracy. It is reasonable to gradually
419 increase the size until this desired accuracy is achieved, to avoid generating a large amount of
420 unnecessary data.

421 ***4.3 Effect of amount of observation wells***

422 Number of observed wells is an important consideration in design a monitoring plan, as it
423 determines measurement efforts and accuracy of any interpretation technique. A dense observation
424 network engages a better understanding of the field when it better covers the dynamics subsurface.
425 However, it also imposes higher requirement on the costs and feasibility of a monitoring plan in a real
426 field sometimes. In this part we examine the importance of measurement points for reconstruction
427 accuracy.

428 To achieve this, the study evaluates two additional configurations of observation wells which share
429 an identical pump setting but using less wells compared to the original configuration containing 49
430 piezometers. The number of wells is first reduced by half (25 wells) and then by one third (16 wells);
431 but we retain the pumping schemes with similar flow rates, paired pumping, and other conditions as
432 performed for the original configuration of 49 wells in Section 3.1 (Figure 8). In practice, we rebuild
433 new datasets in which the hydraulic head maps are interpolated with fewer measurements, which
434 lowers the input resolution and worsens the inversion quality on the 10,000 test models (Table 1).



435

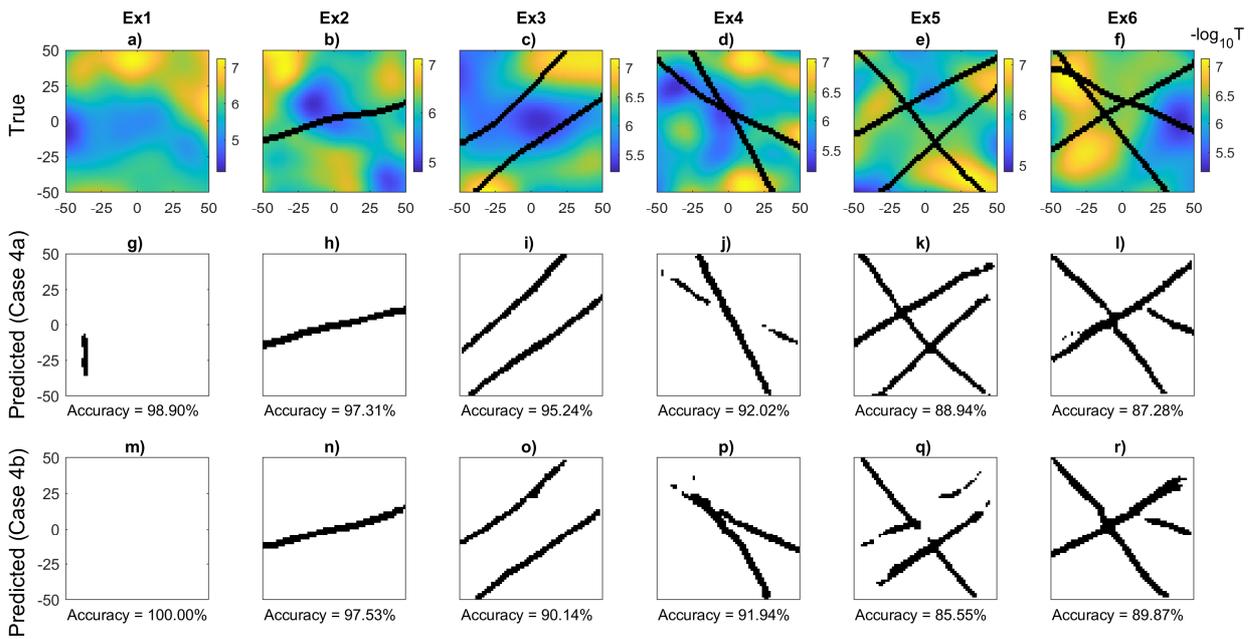
436

Figure 8: Altered monitoring configuration: lowering the number of observation wells to 25

437

while maintaining all other settings: a) Equal spacing, b) Varied spacing.

438



439

440

Figure 9: Reconstructions with a monitoring network of 25 wells (half number of the original

441

design) (Case 4). From top to bottom by row: the truth, prediction, and accuracy of Case 4a and

442

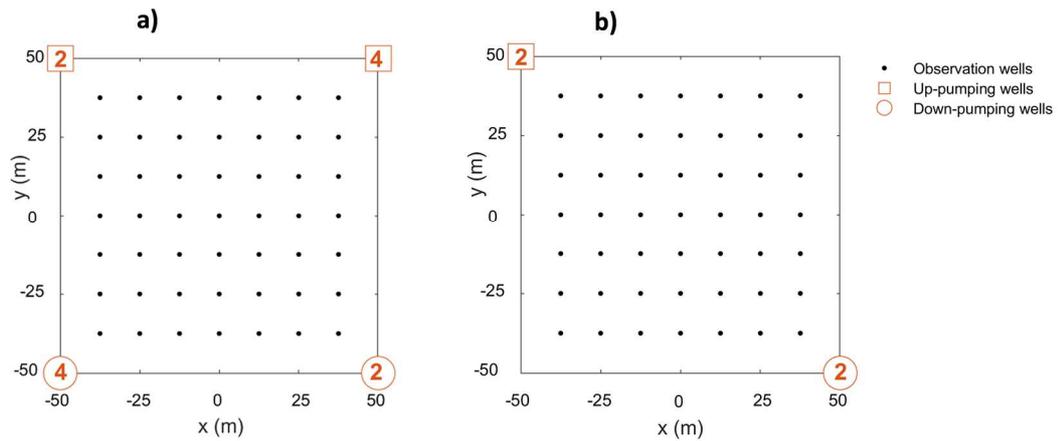
Case 4b; from left to right by column: the representative cases from Ex1 to Ex6. A coarse

443 **observation degrades in the prediction accuracy as the coverage of measurement data is crucial**
444 **in determining the inversion quality.**

445 The results of the 10,000 test models in Table 1 reveal a clear dependence of the reconstruction
446 accuracy on the piezometric density over three configurations. Degraded data coverage results in a
447 drop in overall reconstruction accuracy: from 94.83% (49 wells) to nearly 92% with 25 and 16 wells.
448 The degradation tendency is also confirmed visually in the representative models, where the
449 reconstructions of the fracture network geometry from 25-well setup are less accurate than in the 49-
450 well (Figure 9). For simple fracture networks in Ex1 to Ex3, the reconstruction can be accomplished
451 accurately with only a limited number of observation wells. However, the dissimilarity in the results
452 for 25- and 16-well setups implies that the hydraulic information in these two cases is only sufficient
453 to infer the existence and overall shape of the fractures, but not to identify them more precisely in
454 complex cases. Another comparison of results between two configurations of observation wells in
455 Figure 9 highlights the impact of well spacing. A scattered distribution of observation wells leads to a
456 drop in accuracy of the fracture network reconstruction, as Case 4b shown in Figure 9, where the
457 reconstruction quality depends locally on the scarcity of observation wells. Indeed, the relationship
458 between the number of involved measurements and the reliability of resolved predictions is a common
459 issue in any inversion problem, regardless of the optimization method employed. To ensure an
460 accurate reconstruction, there must be a sufficient number of wells to cover all heterogeneities in the
461 target field.

462 ***4.4 Effect of pumping schemes***

463 To further investigate the impact of the amount of data in identifying fractures, in this section we
464 reduce the number of pumping wells and keep the original configuration with the same pumping rate
465 and observation wells (see Figure 10). This modification in the monitoring scheme results in a
466 reduction of the amount of data to half and quarter of the original scheme.



467

468

Figure 10: Monitoring schemes with less pumping wells: a) 2 pairs of pumping wells, b) 1

469

pairs of pumping wells.

470

The reduction in data size leads to an overall deterioration in the result for the subsurface fracture

471

network reconstruction, with the average accuracy dropping to 93.53% in Case 5a with 2 pairs of

472

pumping wells and 91.50% in Case 5b with a single pair. Details for the six representative examples

473

are shown in Figure 11 while the statistics are summarized in Table 1. Downsizing the observation

474

data in this case undoubtedly leads to less coverage of observed data, resulting in a less reliable

475

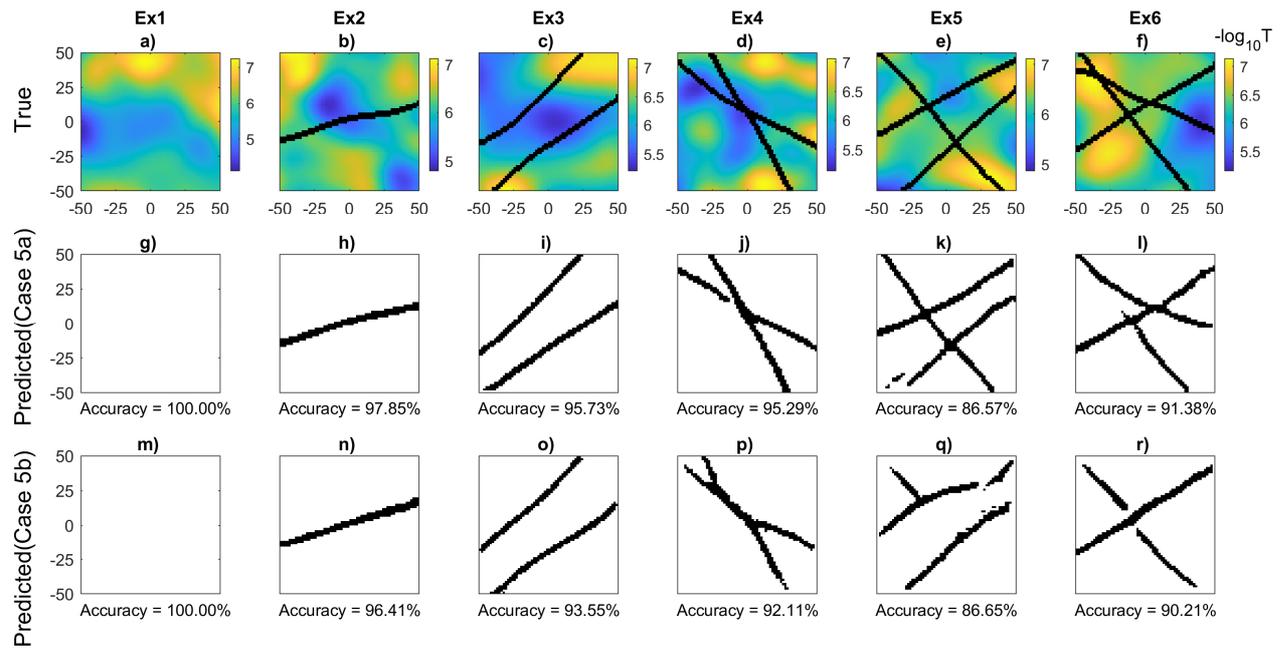
prediction. However, the consequence is less severe than the reduction in observation wells, where the

476

lack of local information results in a coarse resolution that affects details of the reconstructed map.

477

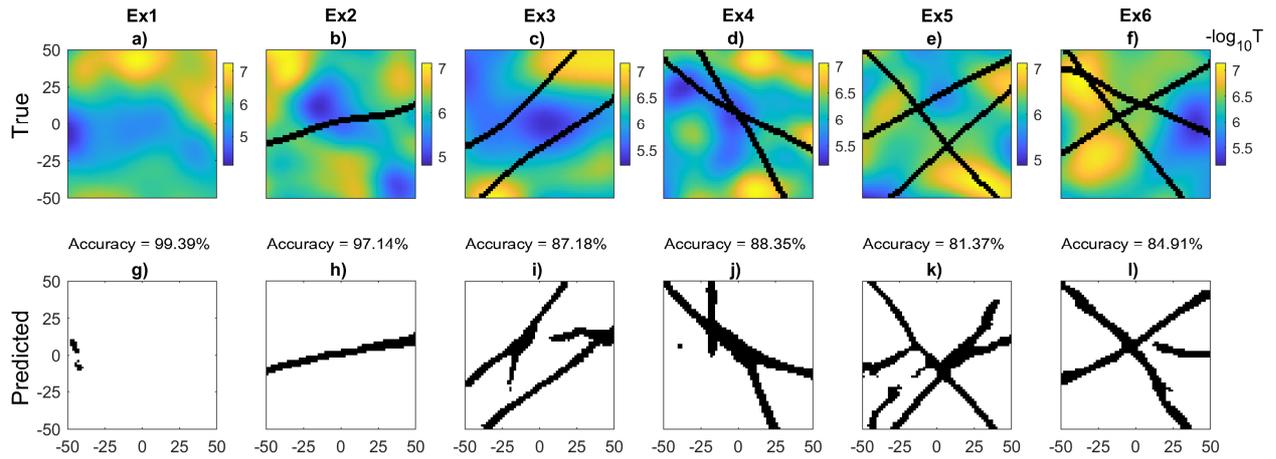
This comparison can be used as a guide to the design of the monitoring system in a real field.



478

479 **Figure 11: Reconstructions with data from different number of pumping wells (Case 5). From**
 480 **top to bottom by row: the truth, prediction, and accuracy of Case 5a (2 pairs of pumping wells)**
 481 **and Case 5b (1 pair of pumping wells); from left to right by column: the representative cases**
 482 **from Ex1 to Ex6. Less observation data degrades the prediction accuracy.**

483 In this section, we analyze the quality of reconstructions with piezometric data obtained with a low
 484 pumping rate of 0.5 l/min while maintaining all other acquisition parameters, such as the number of
 485 pumping wells and observation wells. The low flow rate hence entails a small hydraulic disturbance
 486 confined to the vicinity of the pumping wells. The absence or low response that can be affected by
 487 noise in distant wells cause a lack of information to perform a correct interpretation. Such difficulty
 488 can be the origin of inaccurate reconstructions with artifacts where the interpretation cannot
 489 distinguish the presence of a fracture network from the impacts of the heterogeneity of the
 490 surrounding matrix (see Figure 12).

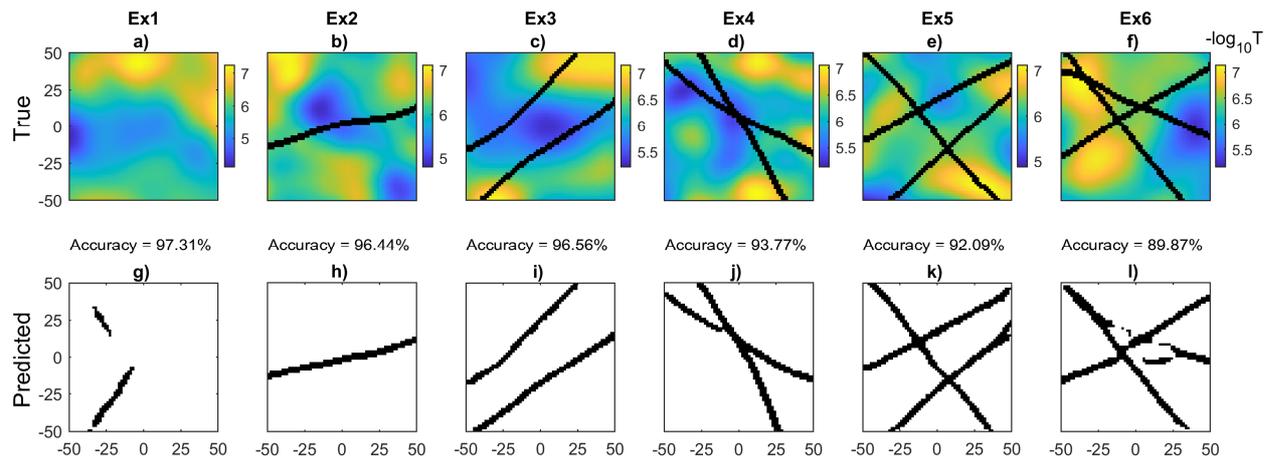


491

492 **Figure 12: Reconstructions with data from a monitoring scheme using a low pumping rate at**
 493 **all pumping wells (Case 5c). From top to bottom by row: the truth, accuracy, and prediction of**
 494 **Case 5c; from left to right by column: the representative cases from Ex1 to Ex6. A low pumping**
 495 **rate results in a small hydraulic disturbance focusing on the vicinity of the pumping wells,**
 496 **which limits the accuracy of fracture detection.**

497 **4.5 Effect of observation uncertainty**

498 Removing the noise that affects hydraulic data can be a complex task to perform prior to
 499 interpreting input data in an inversion algorithm. Thus, identifying the influence of noise on imagery
 500 quality is a crucial step in determining the degree of uncertainty in the interpretation. To analyze this
 501 influence, we contaminate the hydraulic data with random Gaussian noise, where three standard
 502 deviations represent 5%, 15%, and 25% of the original signal, respectively. Contaminated data is then
 503 interpreted using the network formed with the original data (Section 3.3). To illustrate the result for
 504 10000 testing samples, we group the histogram in Figure 11 and summarizes the metric detail in Table
 505 1 while the predictions for the benchmark cases are shown in Figure 10.



506

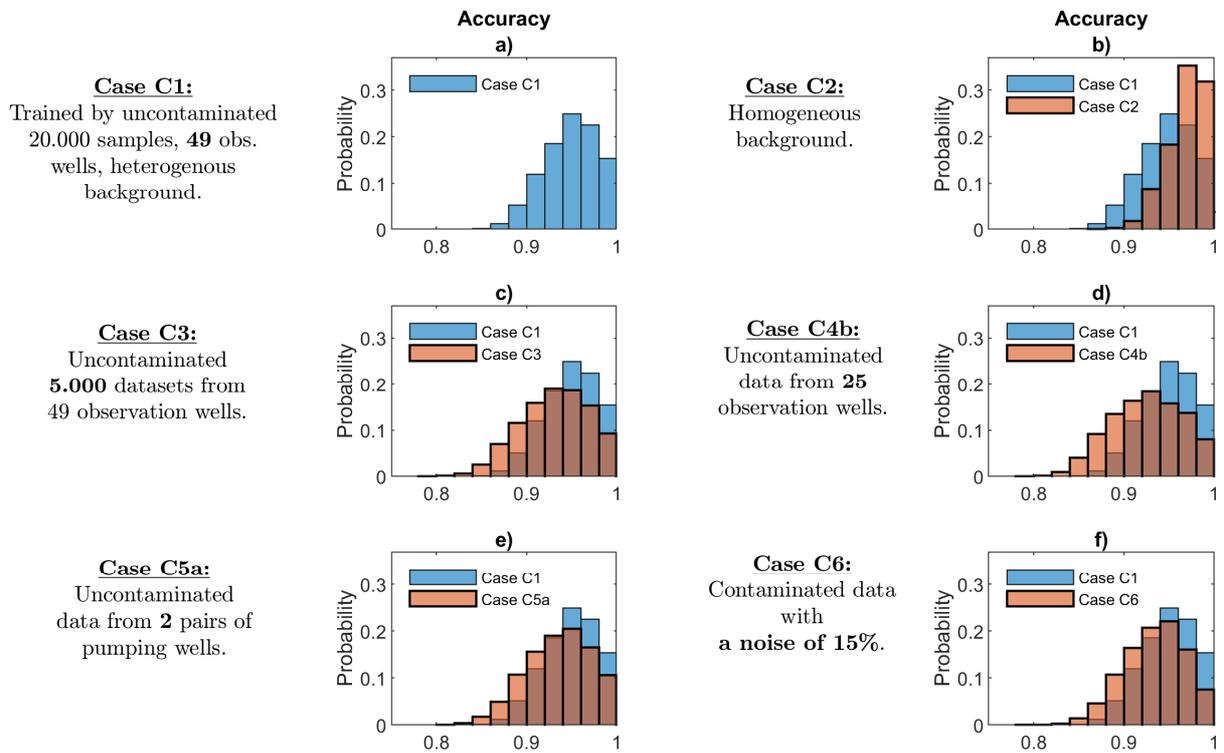
507 **Figure 13: Prediction from measurement contaminated with a Gaussian noise of 15% (Case**
 508 **6). From top to bottom by row: the truth, accuracy, and predictions; from left to right by**
 509 **column: the representative cases from Ex1 to Ex6. The network faithfully reconstructs the**
 510 **fractures network with modest artifacts. Minor noise impact is associated to the neural layers**
 511 **which operates a region-wise convolution and de-noises in processing feature images.**

512 According to the metric assessment (average of all reconstruction accuracies), the increase in noise
 513 correspond to a slight decrease in reconstruction accuracy from 94.59 percent for 5% noise to 93.38
 514 percent for 15% and 91.90 percent for 25% noise. Noise appears to have only a minimal effect on the
 515 reconstruction accuracy of the fractures. Figure 10 details reconstructions of typical models where the
 516 fractures are well reconstructed despite some minor artifacts. Local noise that mimics the dynamic
 517 behavior of fractures in groundwater likely leads to the misinterpretation that produces these artifacts.
 518 Such a disturbance is then attributed to false segments, as in Ex1, or to slight misalignment of
 519 fractures, as in the other cases Ex4 to Ex6. However, the fracture networks are well identified overall,
 520 regardless certain degree of uncertainty in the hydraulic data.

521 This minimal noise effect has been reported in the literature when applying CNN architectures to
 522 solve inversion problems, e.g., in electrical resistivity tomography (Vu & Jardani, 2021), seismology
 523 (Wu & Lin, 2018), or hydrology (Jardani et al., 2022). As explained in Section 2.3, the operation of
 524 CNN network relies on a convolutional mechanism that interpret feature on a region-by-region basis
 525 rather than treating each observation point separately, as is the case with a conventional inversion

526 method. Such a mechanism minimizes the impact of local changes due to random noise. Since the
 527 learning process includes max-pooling operations, which also contribute to denoising the information
 528 by keeping only the maximum values during the process.

529 To summarize, we collect all the result discussed previously in Figure 11 for the histograms and in
 530 Table 1 for the accuracy of reconstructed fractures.



531

532 **Figure 14: Histogram of fracture reconstruction accuracy of 10,000 models in the testing**
 533 **dataset. From top to bottom by row: the cases from Case 1 to Case 6; from left to right by**
 534 **column: contexts of data used in the training and the corresponding histograms. Impacts of**
 535 **interferences is not equal, the accuracy is most sensitive to the number of observed wells, while**
 536 **showing certain resistance to the noise in measurements and heterogeneity in the ground**
 537 **matrix.**

538

539 **Table 1: Summary of prediction accuracy in the discussion above, including the**
540 **representative showcases and the mean for 10,000 models in the testing dataset.**

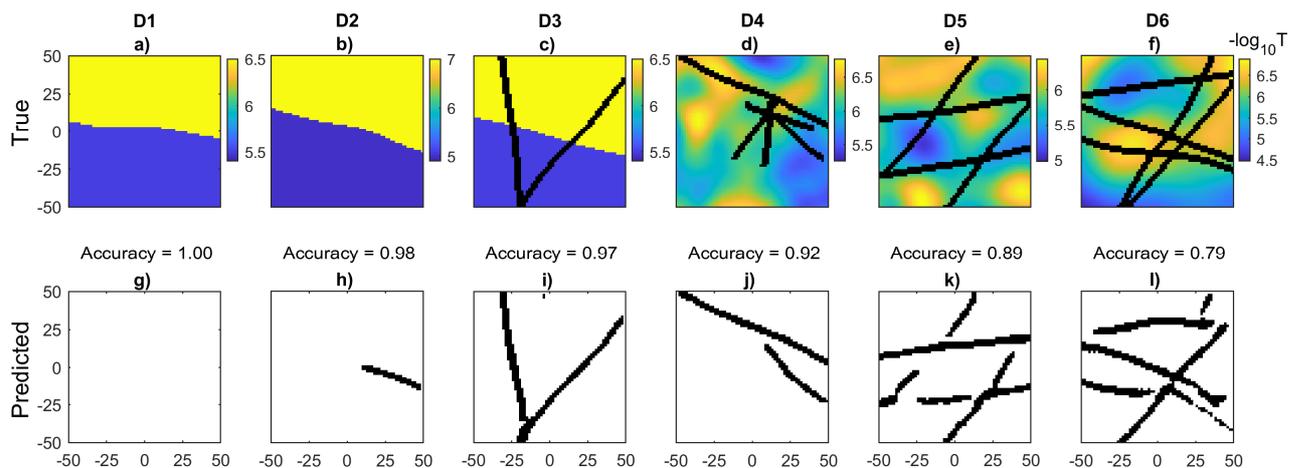
<i>Data type</i>		<i>Case</i>	<i>Accuracy</i>						
			<i>Ex1</i>	<i>Ex2</i>	<i>Ex3</i>	<i>Ex4</i>	<i>Ex5</i>	<i>Ex6</i>	<i>Mean</i>
Complete dataset	20.000 samples, 49 obs. wells, heterogenous ground	C1	100.00	97.31	96.51	94.78	91.94	91.21	94.83
Homogeneous ground		C2	100.00	99.24	98.12	95.95	95.36	95.63	96.94
Size of dataset in training	5.000 samples	C3	100.00	96.44	92.82	93.02	90.06	89.23	93.08
	10.000 samples		100.00	97.44	95.48	91.75	93.68	92.46	93.95
	30.000 samples		100.00	98.32	94.80	93.82	93.14	93.58	95.19
Lower number of observation wells	25 obs. wells	C4a	98.90	97.31	95.24	92.02	88.94	87.28	92.92
	25 obs. wells	C4b	100.00	97.53	90.14	91.94	85.55	89.87	92.53
	16 obs. wells		100.00	97.66	90.31	88.50	84.25	90.63	91.45
Lower number of pump wells	2 pairs	C5a	100.00	97.85	95.73	95.29	86.57	91.38	93.53
	1 pair	C5b	100.00	96.41	93.55	92.11	86.65	90.21	91.50
Lower pumping rate		C5c	99.39	97.14	87.18	88.35	81.37	84.91	90.25
Contaminated with a Gaussian noise	5%	C6	100.00	97.17	95.83	95.19	91.97	91.41	94.59
	15%		97.31	96.44	96.56	93.77	92.09	89.87	93.38
	25%		98.19	95.17	91.58	92.85	90.21	90.14	91.90

541 **4.6 Influence of features in training models**

542 The training datasets, being the main and most essential data, enable machines to learn the feature
543 and make predictions from the learned features. Not only quality and quantity of training datasets but
544 also its relevance therefore affect the prediction accuracy. All the datasets together must be consistent
545 and relevant to the response expected from the neural network. Gathering a broadly featuring data
546 then better approximates the complexity of a real field, but this often associates to overhead costs and
547 challenges in practice. Often, the training dataset is constrained to be as similar as possible to priori

548 knowledge in the field, which implies possible misinterprets if incomplete and inconsistent data
 549 appears. This discussion investigates how well the trained model performs under different conditions
 550 when the presented map differs from the training knowledge.

551 Six synthetic aquifers with different features are tested as shown in Figure 12. On the first three
 552 aquifers (D1, D2 and D3), the transmissivity fields contain binary distributions that differ from the
 553 Gaussian models predefined in the training dataset. The neural network reconstructions successfully
 554 capture the heterogeneities of the aquifer, including the absence or presence of fractures in the target
 555 models. However, in the second case D2, Figure 12b&h, an anomaly is evident due to the
 556 interpretation of a segment at the interface between two hydro-facies. A transmissivity gap at the
 557 hydro-facies interface causes fracture-like behavior in the groundwater, resulting in an artifact in the
 558 prediction. As mentioned in Section 4.1 for Gaussian models, the ground heterogeneity can trigger
 559 misinterpretations depending on the gap of heterogeneity in the ground matrix.



560

561 **Figure 15: Applying trained network to prediction models of different features to the training**
 562 **dataset. From top to bottom by row: the truth, accuracy, and predictions; from left to right by**
 563 **column: the tested cases from D1 to D6. While examples D1 to D3 involve matrix blocks with**
 564 **non-Gaussian transmissivity distribution, the left three examples, D4 to D6, consist of more**
 565 **fracture features. With minimal misinterpretations, the trained network properly reconstructs**
 566 **the fracture systems.**

567 For the remain cases (D4, D5 and D6), we retain the Gaussian nature of the transmissivity as used
568 in the training data, but the fracture network structures are more complicated. In the case D4, the
569 fracture network consists of some short segment which differs from the training data containing only
570 long fractures. The reconstruction of Case D4 efficiently detects the major fracture, but the neural
571 network fails to detect minor fractures that were not learned during the training phase. In general, it is
572 more complex to capture short fractures in hydraulic tomography with a poor resolution of the
573 piezometric coverage (Fischer et al., 2018).

574 In the following cases, D5 and D6, the trained network is challenged to interpret the models with 4
575 and 5 fractures, respectively, which include more features than the trained data with at most 3
576 fractures. The reconstructed maps mostly locate the fracture traces from the interpreted segments.
577 When the complexity of the tested fracture system exceeds the coverage limit of the input data, certain
578 fracture geometries can be simplified. In practice, the nature of fracture networks and the degree of
579 heterogeneity in the transmissivity field covered in the training data reconstruction must be pre-
580 specified based on prior knowledge of the study area. As the rest of the classical deterministic or
581 stochastic inversion methods, inversion with deep learning tools requires priori conditions to constrain
582 the inversion and reduce the non-uniqueness issue in the solution.

583 In this study, the focus is on the theoretical development of an advanced neural network to map the
584 fracture system. Alternative approaches with comparable configuration can be found in the literature
585 (Wang et al., 2016; Wang et al., 2017; Fischer et al., 2018; Ringel et al., 2019) as some of them are
586 addressed and relatively analyzed along the discussions; however, the implementation of a
587 conventional inversion approach is beyond the scope of this work. When processing a real dataset, a
588 more detailed comparison of the approaches should be considered.

589 **Conclusion**

590 In this paper, we present a novel and practical method for identifying fracture systems associated to
591 spatial measurements of the water table in a fractured aquifer. Based on Segnet, the involving
592 advanced neural network is designed to be efficient in terms of memory and processing time during

593 inference, with a substantially lower number of trainable parameters than competing designs. The
594 architecture as a deep fully convolutional neural network, is topologically composed of encoder-
595 decoder structures that directly translate the inverted function in terms of trainable parameters, rather
596 than the indirect approach used in classical inversion techniques.

597 As with any deep learning approach, the algorithm requires a large synthetic training database to
598 establish a reliable generalization capable of predicting the data not seen in the learning stage. The
599 construction of this database was achieved by the geostatistical generation of fracture network models
600 whose number varies between 0 and 3 and whose transmissivity is assumed constant and placed on a
601 rock with heterogeneous matrix transmissivity, also generated according to a geostatistical variogram.
602 Based on these synthetic aquifers, we performed pumping tests according to the hydraulic
603 tomography model by solving the flow equation with the discrete fracture network parameterization.
604 The neural network was trained using 20,000 synthetic aquifers with their hydraulic responses, while
605 another 10,000 datasets test the relevance of the approach. Results show that the trained network
606 succeed in accurately mapping the fracture network geometry and the performance of the network is
607 then discussed in the context of a variety of potential interferences in practice.

608 First, accuracy shows a clear relationship with the size of training dataset. When the network is
609 trained with a limited dataset, the quality of identification of complex fracture networks degrades,
610 while data abundance guarantees a high-quality response. The number of monitoring also affects the
611 quality of the reconstruction, with accuracy improving if the complexity of groundwater dynamics in
612 the fractured aquifer can be captured by the availability of planned monitoring wells. This dependence
613 of inversion results on the number of wells also occurs in conventional deterministic or stochastic
614 inversion methods. However, the neural network shows less impact of data noise on the inversion
615 accuracy, which is due to the convolutional operation and max-pooling that de-noise input data
616 through region-wise interpretation.

617 The quality of the inversion results also depends on the nature of models used for training. As soon
618 as the network is confronted with the processing of hydraulic data from models whose properties are

619 very different from those used in the training, the quality of the inversion deteriorates. The choice of
620 training model features must be based on a priori information in order to obtain an accurate inversion.
621 Like other conventional inversion methods, inversion using a deep learning approach also requires the
622 use of a priori information; and the trained neural network only covers a specific set of models with
623 specific predefined features.

624 A series of tests also reveals the importance of the matrix heterogeneity in reconstructing the
625 fracture network structure. If the transmissivity of the matrix is very low compared to that of the
626 fractures, the matrix contributes little to the dynamics of the flow field and cannot mask the effects of
627 the main fractures. On the other hand, if the matrix contains secondary fractures that are represented
628 together with the matrix in an equivalent porous medium with a slightly permeable transmissivity, this
629 can show the effect of masking the identification of the main fractures.

630 In this work, our effort focuses on the geometric identification of the main fractures of an aquifer
631 where only constant aperture of fractures is considered. However, the fracture characteristics can be
632 more complex bringing more feature such as the hydraulic transmissivity variation along the fracture
633 and between the individual fracture. In the next work, other fracture apertures and the matrix
634 heterogeneity will be tackled which requires the development of a new multi-task network.

635 In this work, our effort focuses on the geometric identification of the major fractures of an aquifer
636 where solely constant aperture fracture is considered. However, in practice fracture characteristics
637 may be more complicated evolving additional features, such as hydraulic transmissivity varying along
638 the fracture and between individual elements of the fracture network. To address this issue, the
639 fracture apertures and matrix heterogeneity are both tackled in the following study, which requires the
640 development of a new multi-task network.

641 **References**

642 Advani M.S., Saxe A.M., and Sompolinsky H., 2020. High-dimensional dynamics of generalization
643 error in neural networks. *Neural Networks*, 132, 428-446.

644 Amari S., Murata N., Muller K.-R., Finke M., and Yang, H.H., 1997. Asymptotic statistical theory
645 of overtraining and cross-validation. *IEEE Transactions on Neural Networks*, 8(5), 985-996.

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

649 Aziz A. R. A. and Wong K.-F. V., 1992. A Neural-Network Approach to the Determination of
650 Aquifer Parameters. *Groundwater*, 30(2), 164-166.

651 Badrinarayanan V., Kendall A., and Cipolla R., 2017. SegNet: A Deep Convolutional Encoder-
652 Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine*
653 *Intelligence*, 39(12), 2481-2495.

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

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

659 Bengio Y., Courville A., and Vincent P., 2013. Representation Learning: A Review and New
660 Perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8), 1798-1828.

661 Benjamin A. T., Elizabeth A. B., and Imme E.-U., 2020. Physically Interpretable Neural Networks
662 for the Geosciences: Applications to Earth System Variability. *Journal of Advances in Modeling*
663 *Earth Systems*, 12(9).

664 Budd D. A. and Vacher H. L., 2004. Matrix permeability of the confined Floridan Aquifer, Florida,
665 USA. *Hydrogeology Journal*, 12(5), 531-549.

666 Dahan O., Nativ R., Adar E. M., Berkowitz B., and Ronen Z., 1999. Field observation of flow in a
667 fracture intersecting unsaturated chalk, *Water Resour. Res.*, 35(11), 3315– 3326.

668 Dausse A., Leonardi V., and Jourde H., 2019. Hydraulic characterization and identification of flow-
669 bearing structures based on multi-scale investigations applied to the Lez karst aquifer. *Journal of*
670 *Hydrology: Regional Studies*, 26, 100627.

671 Dverstorp B., Andersson J., and Nordqvist W., 1992. Discrete fracture network interpretation of
672 field tracer migration in sparsely fractured rock. *Water Resour. Res.*, 28(9), 2327–2343.

673 Elanayar V.T. S. and Shin Y., 1994. Radial basis function neural network for approximation and
674 estimation of nonlinear stochastic dynamic systems. *IEEE Transactions on Neural Networks*, 5(4),
675 594-603.

676 Fischer P., Jardani A., and Lecoq N., 2018. Hydraulic tomography of discrete networks of conduits
677 and fractures in a karstic aquifer by using a deterministic inversion algorithm. *Advances in Water*
678 *Resources*, 112, 83-94.

679 Fukushima K., 1988. Neocognitron: A hierarchical neural network capable of visual pattern
680 recognition. *Neural Networks*, 1(2), 119-130.

681 Ghafoorian M., Karssemeijer N., Heskes T., Uden I.V., Sánchez C., Litjens G., Leeuw F.,
682 Ginneken B., Marchiori E., and Platel B., 2017. Location Sensitive Deep Convolutional Neural
683 Networks for Segmentation of White Matter Hyperintensities. *Scientific Reports*, 7(1), 5110.

684 Illman W. A., 2014. Hydraulic Tomography Offers Improved Imaging of Heterogeneity in
685 Fractured Rocks. *Groundwater*, 52(5), 659-684.

686 Indiveri G. et al., 2011. Neuromorphic Silicon Neuron Circuits. *Frontiers in Neuroscience*, 5, 73.

687 Jardani A., Vu M.T., and Fischer P., 2022. Use of Convolutional Neural Networks with encoder-
688 decoder structure for predicting the inverse operator in hydraulic tomography: CNN-HT. *Journal of*
689 *Hydrology*, 604, 127233.

690 Jiang X., Yu H., and Lv S., 2020. An Image Segmentation Algorithm Based on a Local Region
691 Conditional Random Field Model. *International Journal of Communications, Network and System*
692 *Sciences*, 13, 139-159.

693 Kendall A., Badrinarayanan V., and Cipolla R., 2017. *Bayesian SegNet: Model Uncertainty in*
694 *Deep Convolutional Encoder-Decoder Architectures for Scene Understanding*. Proceedings of the
695 British Machine Vision Conference (BMVC).

696 Khryashchev V., Ivanovsky L., Pavlov V., Ostrovskaya A., and Rubtsov A., 2018. *Comparison of*
697 *Different Convolutional Neural Network Architectures for Satellite Image Segmentation*. 2018 23rd
698 Conference of Open Innovations Association (FRUCT).

699 Kohl T., Evans K. F., Hopkirk R. J., Jung R., and Rybach L., 1997. Observation and simulation of
700 non-Darcian flow transients in fractured rock. *Water Resources Research*, 33(3), 407-418.

701 Laloy E., Héroult R., Lee J., Jacques D., and Linde N., 2017. Inversion using a new low-
702 dimensional representation of complex binary geological media based on a deep neural network.
703 *Advances in Water Resources*, 110, 387-405.

704 Laloy E., Héroult R., Jacques D., and Linde N., 2018. Training-Image Based Geostatistical
705 Inversion Using a Spatial Generative Adversarial Neural Network. *Water Resources Research*, 54(1),
706 381-406.

707 Lary J.A., Alavi H.A., Gandomi H.A., and Walker L.A., 2016. Machine learning in geosciences
708 and remote sensing. *Geoscience Frontiers*, 7(1), 3-10.

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

711 Liu B., Guo Q., Li S., Liu B., Ren Y., Pang Y. Guo X., Liu L., and Jiang P., 2020. Deep Learning
712 Inversion of Electrical Resistivity Data. *IEEE Transactions on Geoscience and Remote Sensing*,
713 58(8), 5715-5728.

714 Maréchal J. C., Dewandel, B., and Subrahmanyam K., 2004. Use of hydraulic tests at different
715 scales to characterize fracture network properties in the weathered-fractured layer of a hard rock
716 aquifer. *Water Resources Research*, 40(11).

717 Mohammadi Z., and Illman W. A., 2019. Detection of karst conduit patterns via hydraulic
718 tomography: A synthetic inverse modeling study. *Journal of Hydrology*, 572, 131-147.

719 Mohammed C. Y. and Edward K., 2019. Semantic segmentation on small datasets of satellite
720 images using convolutional neural networks. *Journal of Applied Remote Sensing*, 13(4), 1-16.

721 Moseley B. and Krischer L., 2020. Chapter One - 70 years of machine learning in geoscience in
722 review. *Advances in Geophysics*, 61, 1-55.

723 Mukhopadhyay P. and Mallick S., 2019. *Bayesian deep learning for seismic facies classification*
724 *and its uncertainty estimation*. SEG Technical Program Expanded Abstracts 2019.

725 Nagi J., Ducatelle F., Di Caro G., Ciresan D., Meier U., Giusti A., Nagi F., Schmidhuber J., and
726 Gambardella L.M., 2011. *Max-pooling convolutional neural networks for vision-based hand gesture*
727 *recognition*. 2011 IEEE International Conference on Signal and Image Processing Applications,
728 ICSIPA 2011.

729 Neuman S. P., 2005. Trends, prospects and challenges in quantifying flow and transport through
730 fractured rocks. *Hydrogeology Journal*, 13(1), 124-147.

731 Park M. J. and Sacchi M. D., 2020. Automatic velocity analysis using convolutional neural network
732 and transfer learning. *Geophysics*, 85(1), 33-43.

733 Pham N., Fomel S. and Dunlap D., 2018. *Automatic channel detection using deep learning*. SEG
734 International Exposition and Annual Meeting.

735 Puzyrev V., 2019. Deep learning electromagnetic inversion with convolutional neural networks.
736 *Geophysical Journal International*, 218(2), 817-832.

737 Puzyrev V. and Swidinsky A., 2021. Inversion of 1D frequency- and time-domain electromagnetic
738 data with convolutional neural networks. *Computers & Geosciences*, 149, 104681.

739 Qin Z., Jiang H, Dai Q., Yue Y., Chen L., and Wang Q., 2020. Robust Lane Detection From
740 Continuous Driving Scenes Using Deep Neural Networks. *IEEE Transactions on Vehicular*
741 *Technology*, 69(1), 41-54.

742 Rawat W. and Wang Z., 2017. Deep Convolutional Neural Networks for Image Classification: A
743 Comprehensive Review. *Neural Computation*, 29(9), 2352-2449.

744 Remy N., Boucher A., and Wu J., 2009. *Applied Geostatistics with SGeMS: A User's Guide*.
745 Cambridge: Cambridge University Press.

746 Ringel L. M., Somogyvári M., Jalali M., and Bayer P., 2019. Comparison of Hydraulic and Tracer
747 Tomography for Discrete Fracture Network Inversion. *Geosciences*, 9(6).

748 Sariturk B., Bayram B., Duran Z., and Seker D. Z., 2020. Feature extraction from satellite images
749 using segnet and fully convolutional networks (FCN). *International Journal of Engineering and*
750 *Geosciences*, 5, 138-143.

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

753 Tiedeman C. R. and Barrash W., 2020. Hydraulic Tomography: 3D Hydraulic Conductivity,
754 Fracture Network, and Connectivity in Mudstone. *Groundwater*, 58(2), 238-257.

755 Trichakis I. C., Nikolos I. K., and Karatzas G. P., 2009. Optimal selection of artificial neural
756 network parameters for the prediction of a karstic aquifer's response. *Hydrological Processes*, 23(20),
757 2956-2969.

758 Vu M. and Jardani A., 2022. Mapping of hydraulic transmissivity field from inversion of tracer test
759 data using convolutional neural networks. CNN-2T. *Journal of Hydrology*, 606, 127443.

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

763 Wang X., Jardani A., and Jourde H., 2017. A hybrid inverse method for hydraulic tomography in
764 fractured and karstic media. *Journal of Hydrology*, 551, 29-46.

765 Wang X., Jardani A., Jourde H., Lonergan L., Cosgrove J., Gosselin O., and Massonnat G., 2016.
766 Characterisation of the transmissivity field of a fractured and karstic aquifer, Southern France.
767 *Advances in Water Resources*, 87, 106-121, 2016.

768 Wu Y. and Lin Y., 2018. InversionNet: A Real-Time and Accurate Full Waveform Inversion with
769 CNNs and continuous CRFs. *arXiv:1811.07875*.

770 Yeh W.W.-G., 1986. Review of parameter identification procedures in groundwater hydrology: the
771 inverse problem. *Water Resources Research*, 22, 95-108.

772 Yeh T. and Liu S., 2000. Hydraulic tomography: Development of a new aquifer test method. *Water*
773 *Resources Research*, 36, 2095-2105.

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

776 Zha, Y., Yeh T.-C., Illman W., Tanaka T., Bruines P., Onoe H., and Saegusa H., 2015. What does
777 hydraulic tomography tell us about fractured geological media? A field study and synthetic
778 experiments. *Journal of Hydrology*, 531, 17-30.

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