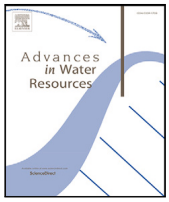
Advances in Water Resources 189 (2024) 104726

Contents lists available at ScienceDirect

Advances in Water Resources

journal homepage: www.elsevier.com/locate/advwatres



A U-Net architecture as a surrogate model combined with a geostatistical spectral algorithm for transient groundwater flow inverse problems

Dany Lauzon

Département des génies civil, géologique et des mines, Polytechnique Montréal, P.O. Box 6079 Station Centre-Ville, H3C 3A7, Montréal, Canada

ARTICLE INFO

Keywords:
Surrogate modeling
Inverse problem
Stochastic hydrogeology
Applied geostatistics
U-Net architecture
Pumping test

ABSTRACT

Characterizing groundwater flow parameters is crucial for understanding complex aquifer systems, and inverse techniques play a fundamental role in modeling hydrogeological parameters and assessing their uncertainties. Nonetheless, the use of a forward model in these methods can be highly time-consuming, especially with an increasing number of model parameters. To address this issue, we propose a surrogate model based on a U-Net architecture that replaces the transient groundwater flow model, reducing runtime and enabling a fast quantification of uncertainties related to key parameters, including heterogeneous hydraulic conductivity, boundary conditions, specific storage, and pumping rate. The surrogate is trained using limited evaluations of the forward model to learn the physical relationship between hydraulic conductivity fields and transient hydraulic heads measured on-site. The physical principles of the studied problem, including boundary conditions, specific storage, and source terms, are also mapped and introduced as inputs to the model to enhance its understanding of the governing equation of transient groundwater flow. To speed up learning using image–image regression, the previously predicted transient hydraulic heads also serve as an input to predict the transient heads at the current time step. Once the model is trained, we use a spectral geostatistical method to solve the inverse problem, a pumping test of 12 h, using the surrogate model in place of the forward model. Our study demonstrates that the trained U-Net accurately reproduces the state variables corresponding to a specific parameter field, and in terms of computational demand, using U-Net as a surrogate model reduces the required computational time by approximately an order of magnitude for the defined problem. The proposed approach offers an efficient and accurate method for groundwater flow parameter characterization and uncertainty quantification in complex aquifer systems.

1. Introduction

Characterization of subsurface heterogeneities represents a crucial initial step in establishing reliable predictions for hydrogeological modeling of groundwater flow (Bárdossy and Hörning, 2015; Taccari et al., 2022). However, a significant challenge arises from the lack of direct observations, which hinders the accurate measurement of specific aquifer properties and gives rise to an inverse problem (Tarantola, 2005; Pasquier and Marcotte, 2006). To address this issue, hydrogeological models must be adapted using inversion techniques and probabilistic approaches based on observed data. Through these methods, unknown parameters of hydrogeological models can be estimated, providing dependable forecasts for more efficient and responsible groundwater resource management (Carrera et al., 2005; Benoit et al., 2017; Laloy et al., 2017; Benoit et al., 2020; Khambhammettu et al., 2020; Bai and Tahmasebi, 2022; Godoy et al., 2022).

Nevertheless, deterministic formulations of inverse problems are known to be inherently difficult to solve due to their ill-posed nature (Dagasan et al., 2020). These formulations often yield non-unique solutions, non-existent solutions, or instabilities. To overcome these limitations, probabilistic methods have emerged as a promising alternative (Mosegaard and Tarantola, 1995; Tarantola, 2005). These approaches allow us to derive probability distributions for parameter values, aligning them with observed state variables, known physical principles, and prior information about the parameters. However, adopting probabilistic techniques introduces significant computational constraints, which can be considered a bottleneck for these methods, as they typically require many iterative executions of the groundwater flow simulator, the forward model (Dagasan et al., 2020; Lauzon and Marcotte, 2022).

A promising approach to reduce computation times is the use of surrogate models (Asher et al., 2015). Rather than opting for full direct modeling, these approaches aim to estimate the response by establishing a correspondence between hydrogeological parameters and state variables. These methods approximate a mapping between the input data (i.e., hydrogeological parameters) and output data (i.e., state

E-mail address: dany.lauzon@polymlt.ca.

https://doi.org/10.1016/j.advwatres.2024.104726
Received 23 August 2023; Received in revised form 20 February 2024; Accepted 17 May 2024
Available online 25 May 2024

variables) necessitating less expensive evaluation in comparison to the underlying physical model. However, it is essential to note that their capability to provide accurate predictions for highly nonlinear mappings with few observational data is limited, and these methods are primarily used in relatively simple hydrogeological systems (Mo et al., 2019a).

Among the promising approaches, the successful application of deep learning methods as surrogate models has been demonstrated in hydrogeological inverse problems in various domains, including petroleum engineering (Tang et al., 2020; Wang et al., 2022), carbon storage (Mo et al., 2019b), contamination source identification (Mo et al., 2019a; Li et al., 2021; Bai and Tahmasebi, 2022; Jiang et al., 2021), steady- and transient-state problems (Tripathy and Bilionis, 2018; Zhu and Zabaras, 2018; Dagasan et al., 2020; Zhang et al., 2022; Taccari et al., 2022; Godoy et al., 2022), subsurface resource modeling (He et al., 2020; Ashworth et al., 2022; Secci et al., 2024), and water resources (Mariethoz and Gómez-Hernández, 2021). However, only a limited number of these approaches have undergone testing to assess their feasibility in capturing the transient behavior of a pumping test, characterized by rapid changes in pressure values, while simultaneously evaluating the uncertainty of hydrogeological parameters. This paper aims to address this gap by evaluating the feasibility and providing a framework to quantify uncertainty in parameters associated with a pumping test, including heterogeneous hydraulic conductivity, boundary conditions, specific storage, and pumping rate.

Several methods have been suggested to substitute the flow simulator with machine learning. We provide an overview of some of these methods that are relevant to the methodology employed here. For instance, Wang et al. (2021) proposed a deep learning framework, a theory-guided neural network, to effectively quantify uncertainty and assimilate flow data to predict hydraulic heads in steady-state problems. They further extended their approach to flow problems in two-phase porous media using theory-guided convolutional neural network (Wang et al., 2022), achieving high accuracy with limited training data for inversion of permeability fields in oil reservoirs. Mo et al. (2019b) developed a substitute based on a deep convolutional encoder–decoder for multiphase flow in geological carbon storage, while Mo et al. (2019a) adopted a similar approach to quantify uncertainty in dynamic multiphase flow for groundwater contaminant source identification in steady state. To account for the temporal relationship between input and output maps, they used the previous timestep's output as input to predict the current timestep. Taccari et al. (2022) proposed the Attention U-Net model to capture physical relationships between inputs and outputs of the aquifer system in steady-state conditions and generate solutions for piezometric levels throughout the domain. This model accurately predicted the steady-state response of a heterogeneous aquifer system by selectively attending to relevant parts of the domain. Tang et al. (2020) used a recurrent surrogate based on the residual U-Net network to assimilate porous oil reservoir data and match the history of a 2D two-phase flow system. Dagasan et al. (2020) explore the use of conditional generative adversarial networks as a surrogate model for hydrogeological inverse problems in a channelized geological structure. The generator is built on a U-Net framework to simulate steady-state pressure head responses to pumping. They combined their surrogate model with the posterior population expansion algorithm (Jäggli et al., 2017) to perform the inversion in a probabilistic manner.

This article proposes a simplified U-Net network as a surrogate model for modeling transient groundwater flow, a 12-hour pumping test, combined with a geostatistical method to determine uncertainties associated with hydrogeological parameters. The U-Net model will learn the correspondence between the inputs and outputs of the model for the two-dimensional equation of transient groundwater flow in a heterogeneous aquifer. The input parameters will provide information on the physical principles of the studied problem, such as boundary conditions, well locations, pumping rate, specific storage, and temporal behavior. Subsequently, a spectral geostatistical method will be used to calibrate complex aquifers to transient conditions. The algorithm used is the S-STBM method, recognized as highly efficient in hydrogeological inverse problems, both for continuous and categorical variable problems (Lauzon and Marcotte, 2020a,b, 2022, 2023). This will enable the characterization of subsurface heterogeneity and the assessment of uncertainty and geological risks associated with groundwater flow in a reduced time. A study of the U-Net network parameters will be carried out to verify the possibility of emulating and replacing the forward operator in a probabilistic inversion process to reduce the computation times associated with inverse problems while maintaining good prediction accuracy.

The rest of the paper is organized as follows. The "Methodology" section introduces the U-Net network architecture and the geostatistical algorithm for the inversion process. The "Numerical Experiment" section presents the groundwater model used, the configuration of the hydrogeological problem, the network training strategy, and performance metrics. In the "Results" section, validation, and testing are performed, and the state of inversions using the U-Net network is compared to inversion using the flow simulator. Specific details about the application of the U-Net network are also provided. The paper concludes with a "Discussion" and a "Conclusion" sections, where a comprehensive discussion occurs.

2. Methodology

The proposed methodology is based on a two-step approach: (1) training a U-Net network to emulate and surrogate the forward model, and (2) resolving the inverse problem using a geostatistical method. The subsequent section elaborates on these two steps.

2.1. Surrogate modeling using U-Net

U-Net was initially proposed for the segmentation of biomedical images (Ronneberger et al., 2015). It is an encoder–decoder network that uses fully convolutional layers and stands out for its ability to work with a very limited number of training samples (Oktay et al., 2018). The encoding process is performed through convolutional layers that progressively reduce the image dimensions while capturing relevant features. Simultaneously, the decoding process uses transposed convolutional layers to restore the original image size while retaining the learned features.

A central aspect of the U-Net model lies in the use of residual connections between the encoder and decoder layers. These connections enable the direct transfer of low-level information from the encoder layers to the corresponding decoder layers, preserving important spatial details during the decoding process. The result is a model capable of achieving robust and accurate performance in various segmentation tasks (Ronneberger et al., 2015; Oktay et al., 2018; Siddique et al., 2021; Wu et al., 2021), and recently in hydrogeological applications (Mo et al., 2019a,b; Dagasan et al., 2020; Jiang et al., 2021; Taccari et al., 2022).

2.1.1. U-Net architecture

The U-Net architecture used in this manuscript is based on the standard U-Net format described by Ronneberger et al. (2015) with slight modifications. The MaxPooling operator has been replaced by convolution layers with strides that halve the block size, following a similar approach to the V-Net architecture (Milletari et al., 2016). Moreover, a single convolution (or transposed convolution) sequence is performed and it is used to downscale (or upscale) the feature maps. The idea behind adopting such a strategy is to reduce the number of trainable parameters to limit memory and speed up training. The model is illustrated in Fig. 1 for 6 input images size of 128×128. Input and output parameters will be explained in Section 3.2.

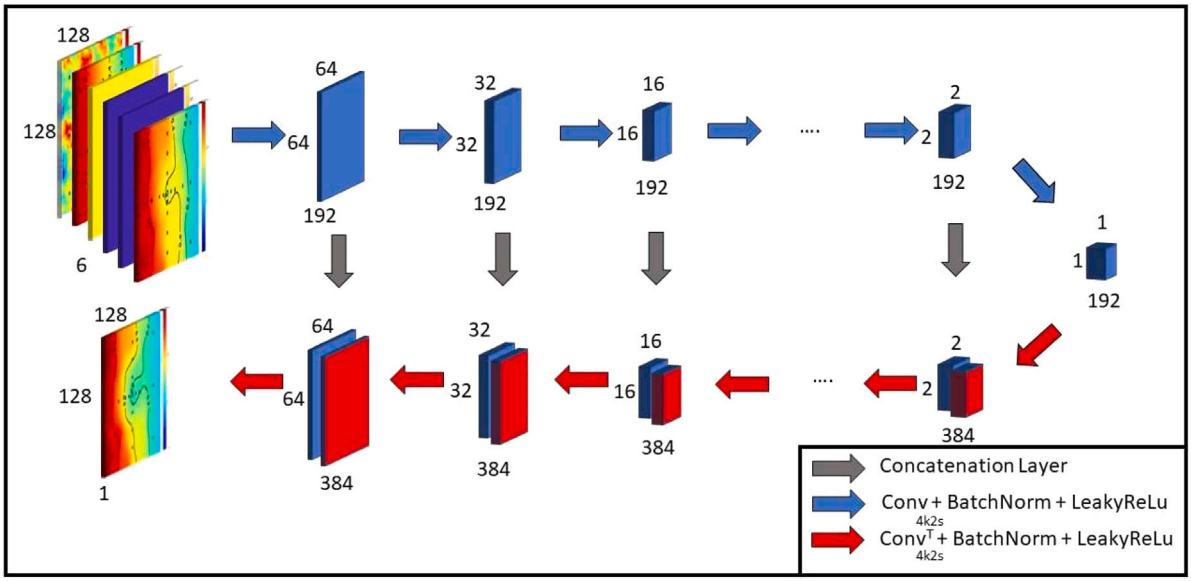


Fig. 1. General overview of the architecture for the modified U-Net with five input maps and one output map.

This modified U-Net architecture has the same symmetrical contraction and expansion networks as the U-Net architecture. Using the architecture presented in Fig. 1, the blue arrows are convolution blocks composed of a 4 × 4 convolution layer, with stride 2 × 2 and padding 'same', a batch normalization layer, and the LeakyRELU activation layer with a 0.2 scale. This step reduces the sampling of the feature map, whose length and width are equal to half the input image. The red arrows represent upscaling blocks consisting of a 4 × 4 transposed convolution layer, with stride 2 × 2 and cropping 'same', a batch normalization layer, and the LeakyRELU activation layer with a 0.2 scale, which increases the scale of the feature map whose length and width are twice those of the input features. The gray arrow represents a copy and concatenate operation whereby the blue feature maps are concatenated with the red feature maps.

At each stage, the number of channel dimensions is maintained at 192 to reduce learning time and the number of learnable variables. This number of channels is achieved by employing a group convolution layer as the first convolution layer with 32 channels per input, resulting in a dimension of 192 channels (32 × 6). This modified U-Net architecture consists of 50 layers with a total of 10.0 million learnable variables. We simplify the network to a minimum to verify its applicability in predicting transient state variables in a heterogeneous domain with a fast training time and reasonable accuracy. Note that the classical U-Net architecture proposes to double the number of feature channels at each down-sampling step, i.e., passing from 32 channels to 64, to 128, and so on. Some researchers stop at 512, sometimes 256, or 1024 (Ronneberger et al., 2015; Dagasan et al., 2020; Jiang et al., 2021; Taccari et al., 2022). If we decide to double the number of feature channels at each downscaling step to reach 512 channels, our model will contain hundreds of millions of learnable parameters, a prohibitive number necessitating high computational demands.

2.1.2. Loss function

The objective of the image–image regression task is to minimize the difference between the predicted image $\hat{\boldsymbol{y}}$ and the target training image $\boldsymbol{y}$. To achieve this, the half-mean squared error (HMSE) loss function is employed, given by the following formula:

$$\text{HMSE Loss} = \frac{1}{2N_b}\sum_{i=1}^{N_b}(\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i)^2 \tag{1}$$

where N_b represents the mini-batch size used during training.

2.2. Inverse problem with S-STBM

This section introduces the probabilistic formulation of the inverse problem following (Tarantola, 2005), using the sequential spectral turning bands method (S-STBM) as the geostatistical approach (Lauzon and Marcotte, 2020a). The main objective of the probabilistic inverse problem is to determine the distribution of aquifer parameters based on a set of observed state variables and their associated measurement errors. These state variables may include hydraulic heads at steady- or transient-state, geochemical concentration transport, travel times between wells, or responses to tracer tests. For our study, transient heads from a 12-hour pumping test act as the state variables.

2.2.1. Inverse problem

According to Tarantola (2005), a model denoted as $m = \{m_1, m_2, \ldots, m_n\}$ represents a finite set of parameters that fully describe the physical system under consideration. For example, it could be a hydraulic conductivity map. To generate these models (or maps), various geostatistical techniques based on the Gaussian hypothesis may be used, like sequential Gaussian simulation (Deutsch, 1992), discrete spectral methods (Dietrich and Newsam, 1993; Chilès and Delfiner, 1997; Le Ravalec et al., 2000) and continuous spectral methods (Shinozuka, 1971; Shinozuka and Jan, 1972; Shinozuka and Deodatis, 1996; Lantuéjoul, 2002; Emery et al., 2016). In our study, models (or maps) are generated using the S-STBM approach, which is renowned for its effectiveness in addressing inverse hydrogeological problems (Lauzon and Marcotte, 2020a,b, 2022, 2023).

Once a model m is defined, the response is calculated using the direct operator g, which solves the partial differential equations describing groundwater flow. This mapping predicts model responses, denoted as $d = g(m)$, at observation points. The calculated values d are then used to evaluate how well a given model m can reproduce the observed data d_{obs}. The evaluation can be quantified by an objective function (OF) that aims to minimize the difference between modeled and observed data, assuming a specific distribution of measurement errors.

In S-STBM, the OF is based on the error between the observed (d^{obs}) and the modeled data ($g(m)$):

$$OF = MSE(m) = \frac{1}{n_{obs}}\sum_{i}^{n_{obs}}[g_i(m) - d_i^{obs}]^2 \tag{2}$$

where n_{obs} represents the number of observation points. A stopping criterion can be implemented to halt the inverse problem when the objective function falls below the threshold associated with the measurement errors.

2.2.2. S-STBM

The S-STBM algorithm is a geostatistical approach that solves inverse problems by calibrating a Gaussian random field using a combination of several cosine functions. The optimization process involves generating independent frequency vectors $\boldsymbol{v}_i$ from the spectral densities associated with the desired covariance function. These vectors are then paired with optimized phases u_i which minimizes an OF. The algorithm is sequential, optimizing each phase u_i in sequence. The Gaussian random field is built from scratch using the following equation:

$$\boldsymbol{Z}_i(\boldsymbol{x}) = \sqrt{\frac{i-1}{i}}\boldsymbol{Z}_{i-1}(\boldsymbol{x}) + \sqrt{\frac{1}{i}}\sqrt{2}cos(\langle \boldsymbol{v}_i, \boldsymbol{x}\rangle + 2\pi u_i) \tag{3}$$

where $\boldsymbol{Z}_i(\boldsymbol{x})$ represents the Gaussian random field at the location of grid node $\boldsymbol{x}$ of iteration i, $\boldsymbol{v}_i$ is an independent frequency vector sampled from the desired spectral density function, $\langle\cdot,\cdot\rangle$ denotes the inner product, and u_i, assumed to be uniformly distributed in $[0, 1]$, is the calibrated phase at iteration i. Typically, at least hundreds of iterations are needed to converge to the multivariate Gaussian distribution (Lantuéjoul, 2002; Lauzon and Marcotte, 2020b).

The underlying assumption of S-STBM is that the distribution of the set $\boldsymbol{u} = \{u_i | i = 1, \ldots, L\}$, L the number of optimized phases, which should follow a uniform distribution in [0,1], is not influenced by the optimization of each phase u_i. However, this assumption is approximated, as the optimized phases may deviate from a uniform distribution. Nevertheless, the case studies presented in Lauzon and Marcotte (2020a, 2022, 2023) suggest that optimization has a negligible impact on the covariance function, as the realization variograms remain close to the target variogram.

The optimization of each phase u_i in S-STBM is achieved using a few golden search steps (typically < 10 steps) (Luenberger and Ye, 2008). To ensure that the Gaussian random field has no visible periodicities, a minimum of $L = 1,000$ cosine functions is generated. For covariance models with linear behavior at the origin, the value of L may need to be increased to ensure sufficient sampling at high frequencies. In this case, $L = 4,000$ cosine functions are used. For more details on the S-STBM method, in particular, how to sample frequency vectors from a desired covariance function, refer to Lauzon and Marcotte (2020a, 2023).

To accelerate the creation of the optimized Gaussian field, we adopt the approach proposed by Lauzon and Marcotte (2023) for phase optimization. Instead of sequentially optimizing one phase at a time, we optimize a batch of m phases. The gradual deformation method (Hu, 2000) is employed to achieve this. We merge two Gaussian white noise vectors of length l, denoted as $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$. The Gaussian values are then transformed into values of the cumulative normal distribution function ($G(\cdot)$), providing a set of phases to be used in Eq. (3):

$$\boldsymbol{u} = G(\boldsymbol{y}_1\cos(t) + \boldsymbol{y}_2\sin(t)) \tag{4}$$

As a result, a single continuous variable, t, defined on $[0, 2\pi]$, needs to be optimized instead of l phases. This substantially reduces the number of calls required by the flow simulator. In this study, we preserve solely those optimized vector phases that result in a reduction of the OF. Conversely, we discard the remaining phases, akin to an accept/reject algorithm.

2.2.3. Simulate single-value property using S-STBM

Given the uncertainties in boundary conditions, pumping rate, and specific storage, it is imperative to devise a method for estimating the posterior distribution of all these parameters, including the hydraulic conductivity field. In response to this necessity, we have adapted S-STBM to calibrate fields representing single-value properties.

Consider simulating a single-value property at a specified coordinate, denoted as $\boldsymbol{x} = x_0$, using S-STBM. S-STBM generates a zero-mean unit variance random field at this coordinate, employing a chosen covariance function (e.g., a spherical model). The specific choice of the covariance function is inconsequential, as we simulate only one coordinate with a Gaussian distribution. We can transition from the normal distribution to the marginal distribution of the property through Gaussian anamorphosis. This graphical process associates Gaussian quantiles with the marginal distribution of the property. Thus, we obtain a value adhering to the prior distribution of the parameters. Subsequently, optimizing the phase linked to this Gaussian field allows the determination of the posterior distribution of the property by solving the inverse problem.

To simultaneously optimize the hydraulic conductivity field, boundary conditions, pumping rate, and specific storage, Eq. (4) can be extended to include the phases associated with the multiple fields generated by S-STBM. Consequently, we need to simulate five fields: the hydraulic conductivity with 128 times 128 parameters, the two boundary conditions represented by two fields of a single parameter, and similarly for the pumping rate and the specific storage. Therefore, we can optimize all five properties using only one parameter in every iteration.

3. Numerical experiment

The objective of this paper is to validate the use of a U-Net architecture as a surrogate model for groundwater flow simulation in transient states, aiming to accelerate the resolution of an inverse problem using a geostatistical method such as S-STBM. The feasibility of U-Net in capturing the transient behavior of a pumping test, characterized by rapid changes in pressure values around a well, is tested. S-STBM is employed to assess the uncertainty of key parameters determined by the pumping test, such as the hydraulic conductivity field and specific storage, while also addressing modeling errors associated with boundary conditions and pumping rates. In this context, reality is assumed to be perfectly known and simplified to rigorously test the quality of the results obtained with the proposed methodology.

The methodology is illustrated in Fig. 2 and is divided into two sections. The first involves training the network using a dataset generated by a flow simulator. The network is trained to recognize highly nonlinear constraints between the inputs (hydraulic conductivity, physical principles, and previous transient heads) and the output (current transient heads at time t). Once the network is trained, it can be used as a surrogate model to solve an inverse problem using the spectral geostatistical method, S-STBM. The inverse problem aims to ascertain the hydraulic conductivity field, boundary conditions, specific storage, pumping rate, and their uncertainties from a collection of groundwater head measurements around a pumping well during a pumping test. Detailed explanations regarding the setup of this experiment are presented in the following subsections.

3.1. Groundwater model

Without loss of generality, consider a two-dimensional groundwater flow in a heterogeneous aquifer. Subject to appropriate boundary conditions, the transient-state groundwater flow in saturated media satisfies the fundamental governing equation:

$$\mathrm{div}(K\nabla h) + Q = S_S\frac{\partial h}{\partial t} \tag{5}$$

where h is the hydraulic head $[L]$, K is the hydraulic conductivity $[LT^{-1}]$, Q is a source term $[T^{-1}]$, S_s is the specific storage $[L^{-1}]$, and t is the time $[T]$.

The MATLAB Reservoir Simulation Toolbox (MRST), a freely available open-source software designed for modeling hydrogeological systems (Lie, 2019), serves as the flow simulator for generating the training images. Note that the MRST is primarily designed as a research tool for quickly prototyping new simulation methods and modeling concepts, rather than serving as a dedicated flow simulator. In this study, transient head data were acquired by solving a non-linear physical model through automatic differentiation on a regular grid.

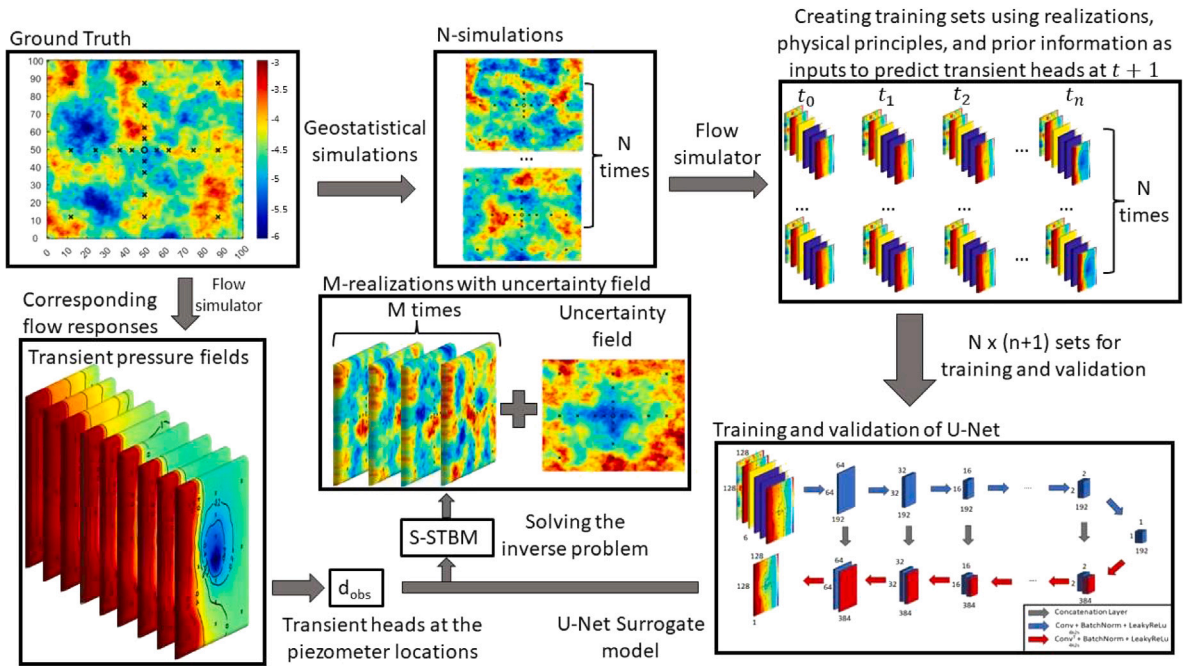


Fig. 2. General overview of the methodology. A Gaussian random field was used to create a reference log-hydraulic conductivity field and its corresponding transient flow responses. Geostatistical simulations were used to create N sets of hydraulic conductivity fields with the desired covariance function using different seed values. N × (n + 1) input sets were created to train the U-Net architecture using an autoregressive strategy. The trained model was then used as a surrogate in the geostatistical inversion method, S-STBM, to generate M realizations and assess uncertainty.

3.2. Problem setting

The synthetic aquifer represents a confined aquifer measuring 100 m × 100 m, discretized on a 128 × 128 regular grid. The depth of the aquifer is assumed to be constant throughout the area, set at 3 m. The upper and lower sides are designated as no-flow boundaries (i.e., Neumann boundary conditions), while the fixed head boundary conditions, namely Dirichlet boundary conditions, are set to 1.042 m and −0.035 m on the left and right sides, respectively.

A pumping well with a constant pumping rate of 17 L/min is located at the center of the area. Transient groundwater flow is simulated with an 18-step schedule ranging from 0 to 12 h. Pressured heads were obtained at 9 various time intervals: 0 min (initial state), 1 min, 5 min, 15 min, 30 min, 60 min, 120 min, 180 min, and 720 min (close to steady state) over 16 piezometers, providing the conditioning data (9 × 20 = 180 head data points). A random Gaussian noise with σ = 0.05 m was added to the extracted head values to simulate measurement errors. The locations of the 20 piezometers (black crosses) were illustrated in Figs. 3 and 4. Hydraulic conductivity follows a base-10 log-normal distribution with a log mean of −4.53. The covariance function used to model the log conductivity is an isotropic spherical model with a range of 30 m and a sill of 0.25. The specific storage, S_S, is set to 8.13E−5 m^{-1}.

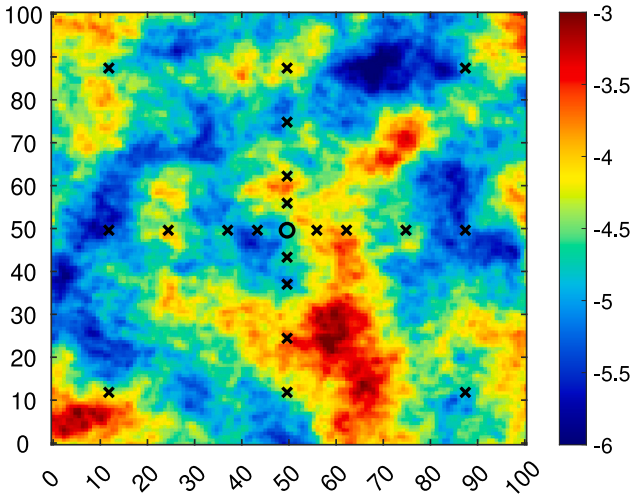


Fig. 3. The reference log-hydraulic conductivity field that was considered the ground truth (Black crosses: locations of piezometers, Black circle: location of the pumping well).

3.3. Network training

The network takes as input an image consisting of six channels: the hydraulic conductivity field, a map summarizing the boundary conditions, two inputs representing the source term, one input summarizing the specific storage value, and the previous prediction of transient hydraulic heads. The output map represents the transient heads at the current time step. Fig. 5 shows the input and output maps for a given time t.

The two Dirichlet boundary conditions can be represented by an image where the values to the east and west of the image are the imposed hydraulic heads. The hydraulic heads throughout the image are subsequently interpolated using a linear function, ensuring that the flow is perpendicular to the north and south boundaries, thus respecting the no-flow conditions imposed by the Neumann boundary conditions (See Fig. 5b). Incorporating boundary conditions into image-to-image regression allows the network to assimilate and understand the physics of the problem.

The source term can be interpreted through two attributes: intensity (or pumping rate) and active time. These two attributes are incorporated into separate maps, constructed by incorporating the pumping rate (See Fig. 5c) and the time since the well activation (See Fig. 5d) within the cell where the source term is situated. All other cells are assigned zero values. At the initial state (t = 0), the intensity of the source term yields a pumping rate of zero.

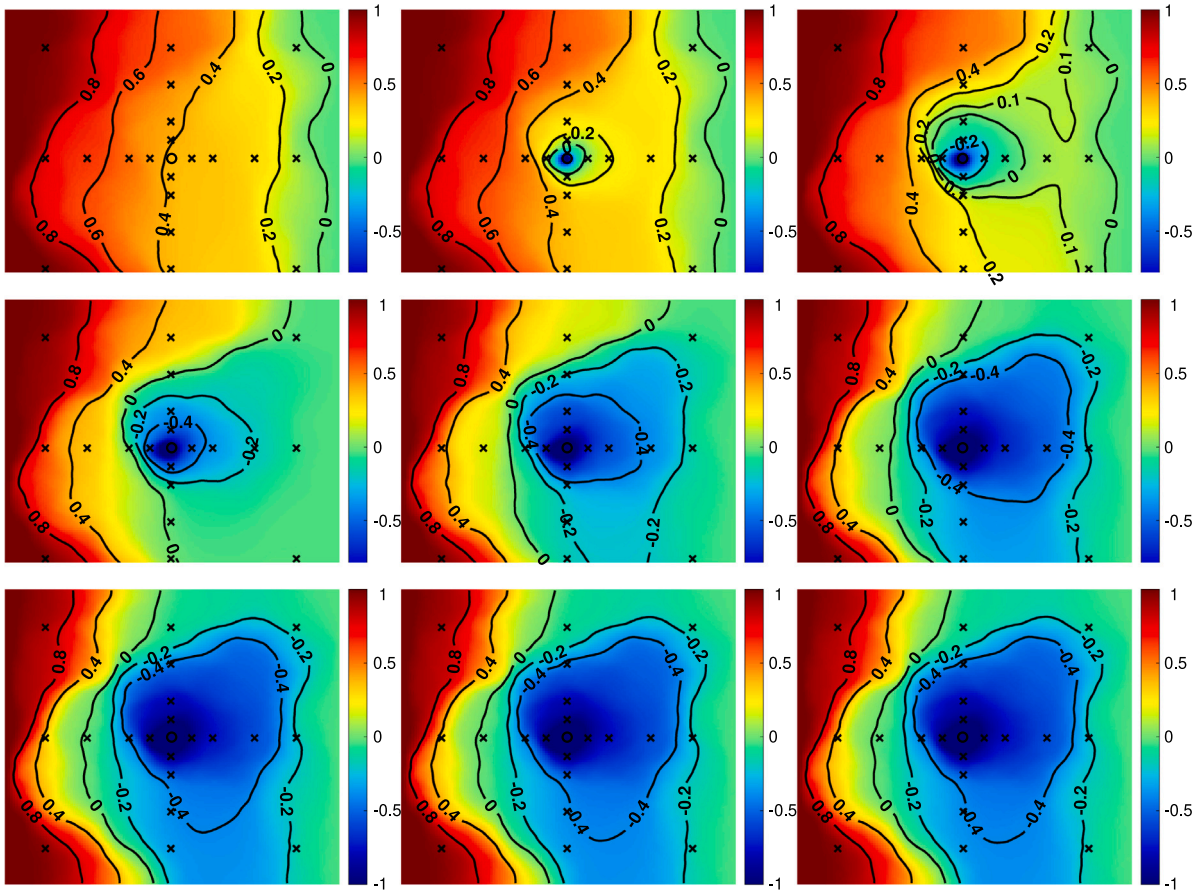


Fig. 4. Corresponding transient flow responses of the ground truth. The corresponding times from the top left to the bottom right are: t = 0, 1, 5, 15, 30, 60, 120, 180, and 720 min. (Black crosses: locations of piezometers, Black circle: location of the pumping well, black lines: contour lines).

The specific storage is also integrated using a map where the specific storage value is placed within the cell where the source term is situated. All other cells are assigned zero values. In our study, we assume the specific storage to be a constant value. However, one can generate a heterogeneous storativity field in the same manner as the hydraulic conductivity field and use the mapping in the training process. For a constant value, our tests yield better results by only providing the value at the well location, with the remaining cells filled with zeros (See Fig. 5f).

In this work, the boundary conditions, specific storage, and pumping rate are treated as uncertain parameters. For training the network, these values were sampled from uniform distributions: [0.9, 1.1] m for the left, and [−0.1, 0.1] m for the right side of the boundary conditions, 10 L/min to 20 L/min for the pumping rate, and log-uniformly between −5 and −3 for the specific storage, corresponding to values between 1E-5 and 1E-3. Sampling values from a uniform distribution during training facilitates exposure to a wide array of potential boundary conditions, specific storage values, and pumping rates.

The model was trained in a supervised manner by producing 600 flow simulations using MRST, each with a distinct hydraulic conductivity field, boundary conditions, specific storage, and pumping rate (i.e., a different seed is used for each simulation). It is worth noting that the system modeling generated a set of 9 output maps, one for each time step (t = 0, 1, 5, 15, 30, 60, 120, 180, and 720 min). Thus, for each flow simulation, it was possible to create 9 sets of input data following the methodology illustrated in Fig. 5 and explained subsequently. The inherently sequential nature of this autoregressive strategy facilitates the acquisition of a substantial dataset for network training and validation purposes.

Therefore, for the 600 flow simulations, 5 400 data sets were created and used for training (3600), validation (900), and testing (900). The validation sets were utilized to fine-tune the hyperparameters of the model, while the testing sets were employed to assess performance metrics. With this strategy, the size of the training dataset was large enough to ensure the model's generalization capability. The losses were minimized using the Adam optimizer (Kingma and Ba, 2014) with an initial learning rate α = 1E−3, and a decreasing rate of 0.8 every 10 epochs. The network was trained for 200 epochs with minibatches of 32. The training converged after approximately 2 h and 54 min, using an NVIDIA GeForce RTX 3070 Ti graphics processor. The U-Net was developed and trained in the MATLAB Deep Network Designer toolbox.

3.4. Performance metrics

The performance of the trained network is assessed by measuring the mean absolute error (MAE), the root mean squared error (RMSE), and the coefficient of determination (R^2) between corresponding pixel values in the target image and the predicted image. These evaluation metrics provide insights into the accuracy and overall quality of the model's predictions compared to the ground truth.

$$\mathrm{MAE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|} \tag{6}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \tag{7}$$

$$\mathrm{R}^2 = 1 - \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{8}$$

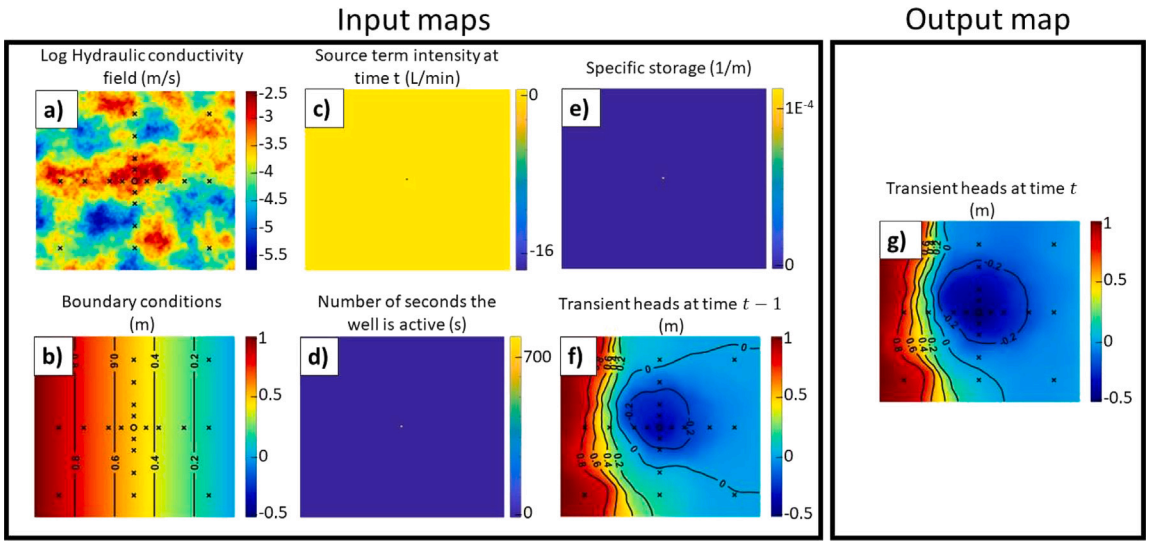


Fig. 5. Overview of input and output maps. U-Net sequentially predicts the model output at time t (g) given an input hydraulic conductivity map (a), an image representing the boundary conditions (b), the mapping of the localization source term with source intensity (c), and the number of seconds during which pumping is activated (d), the map of predicted transient heads at time $t-1$ (e), and the mapping of the specific storage (f).

Table 1
Summary of the four scenarios: If 'Yes', the parameter has variability. If 'No', the parameter is fixed. (Notation: K - Hydraulic conductivity, $B.C.$ - Boundary conditions, S_S - Specific storage, Q - Pumping rate).

Scenario	K	$B.C.$	S_s	Q
1	Yes	No	No	No
2	Yes	Yes	No	No
3	Yes	Yes	Yes	No
4	Yes	Yes	Yes	Yes

where N denote the total number of pixels of each image, y represents the target image, $\hat{y}$ represents the network prediction, $\bar{y}$ is the mean of the target image (y) in the dataset. A well-performing trained model is characterized by minimal MAE and RMSE values approaching zero, accompanied by a R^2 close to one.

4. Results

In this section, we begin by evaluating the network's performance in accurately identifying transient pressure heads during a 12-hour pumping test. We provide a sensitivity analysis regarding the number of model evaluations during training and the number of epochs. Subsequently, we compare the inversion results obtained from S-STBM, utilizing the U-Net network as a surrogate model, with those acquired from S-STBM without surrogate modeling, i.e., using the forward model.

Four scenarios are under study. In the first scenario, we maintain fixed boundary conditions, specific storage, and pumping rates. This means that the network is exclusively trained on diverse hydraulic conductivity fields. In the second scenario, we introduce variations in the boundary conditions. The third scenario involves incorporating different values of specific storage during training. Lastly, the fourth scenario introduces a range of pumping rates to ensure a more comprehensive and diverse training set. These four scenarios are designed to assess the impact of including multiple parameters in the image-to-image regression tasks on the training and validation of the U-Net architecture. Table 1 resumes the four scenarios.

4.1. U-Net validation

In all four scenarios, the same methodology was employed. A total of 400 models were used for training, 100 models for validation, and 100 for testing, resulting in 3600 training sets, 900 validation sets, and 900 testing sets. The distribution of parameters is summarized in Table 2. We trained four networks, one for each scenario listed in Table 1. The objective is to evaluate the impact of adding parameters on the efficiency of the networks.

4.1.1. Testing the network on four scenarios

We showed two models associated with scenario 1 (Fig. 6), and two models associated with scenario 4 (Fig. 7). The hydraulic conductivity field and their corresponding pressure fields at each time step (t = 0, 1, 5, 15, 30, 60, 120, 180, 720 min) predicted by U-Net were visually compared to the maps obtained from the flow simulator. Additionally, scatterplots are displayed, offering a comprehensive view of the U-Net network's performance. The results demonstrate that the generated images are quite similar to their originals, though slightly noisy, as depicted by the contour lines in the U-Net predictions. Also, the main errors are primarily concentrated around the pumping influence zone as depicted by the scatterplots. Moreover, Scenario 4 provides more noisy images than Scenario 1 when comparing visually the contour lines of both scenarios.

The testing process utilized two key metrics: Root Mean Squared Error (RMSE) and the coefficient of determination (R^2). All scenarios provide a training average RMSE below 0.05, which is lower to the errors added by the Gaussian noise of $\sigma = 0.05$ to the observed data, and an average R^2 score higher than 0.98. These results suggest that the trained U-Net model can effectively predict highly nonlinear mappings for the studied case, regardless of the number of uncertain parameters (i.e., the scenario type). However, due to the sequential nature of the approach, where time $t-1$ is used to predict time t, errors from the previous iteration are propagated to the current one. This phenomenon was quantified by an increase in RMSE (the same for MAE, not shown) and a decrease in R^2 score as time progressed. This behavior is depicted in Table 3. Moreover, testing error increases as uncertain parameters are added to the network. This suggests that to obtain similar testing results as in Scenario 1, one may need to add more models in the training sets or increase the number of epochs to train a network dealing with more parameters than in Scenario 1. This would inevitably lead to more computation time and possibly necessitate a readjustment of the hyperparameters. Another way is to redesign the network by adding more channels, resulting in more learnable variables and a more complex training process. We do not explore such approaches, as the results obtained are satisfactory for our purposes.

Table 2
Parameters distribution for the training, validation, and testing sets.(Notation: K - Hydraulic conductivity, $B.C.$ - Boundary conditions, S_S - Specific storage, Q - Pumping rate).

Parameter	Distribution	Value range	Unit
K	Base-10 Log-Normal	Approximately between [1E−6, 1E−3] Isotropic spherical model Log mean: −4.53 Sill: 0.25 Range: 30 m	m/s
$B.C.$ (Left side)	Uniform	[−0.1, −0.1]	m
$B.C.$ (Right side)	Uniform	[0.9, 1.1]	m
S_S	Base-10 Log-Uniform	[−5, −3] (i.e., [1E−5, 1E−3])	m^{-1}
Q	Uniform	[10, 20]	L/min

Table 3
Time evolution of the validation metrics (RMSE and R^2) for 100 realizations resulting in 900 datasets. Predictions are made sequentially, leading to error propagation. This results in an increase in RMSE and a decrease of R^2 over time.

Scenario	1		2		3		4	
Metrics	RMSE (m)	R^2 (–)	RMSE (m)	R^2 (–)	RMSE (m)	R^2 (–)	RMSE (m)	R^2 (–)
t = 0 min	0.014	0.997	0.015	0.997	0.015	0.997	0.017	0.996
t = 1 min	0.020	0.994	0.024	0.992	0.029	0.990	0.035	0.984
t = 5 min	0.027	0.991	0.032	0.988	0.041	0.984	0.048	0.976
t = 15 min	0.034	0.900	0.039	0.985	0.047	0.983	0.049	0.981
t = 30 min	0.038	0.989	0.044	0.984	0.054	0.980	0.051	0.981
t = 60 min	0.042	0.988	0.048	0.983	0.056	0.980	0.053	0.982
t = 120 min	0.044	0.987	0.050	0.982	0.054	0.982	0.054	0.983
t = 180 min	0.045	0.987	0.051	0.982	0.056	0.981	0.058	0.981
t = 720 min	0.046	0.987	0.052	0.981	0.057	0.981	0.061	0.980
Mean	0.034	0.990	0.039	0.986	0.045	0.984	0.047	0.983

4.2. Impact of epochs and training sizes

As is common with many deep learning algorithms, training models requires a significant amount of time. An analysis was conducted by altering the number of epochs and the size of the training dataset while monitoring variations in terms of Mean Absolute Error (MAE), Root Mean Square Error, and R^2, along with training time. The same hardware configuration was used, specifically an NVIDIA GeForce RTX 3070 Ti graphics processor. The experiment is conducted using Scenario 4. Similar results are obtained for Scenarios 1 to 3.

The results of this study, presented in Table 4, provide detailed insights into the MSE, RMSE, R^2, and training time obtained for various combinations of epoch numbers (i.e., 100, 200, 300, 400, 500) and the number of hydrogeological models generated for training (i.e., 100, 200, 300, 400, 500). We recall that a hydrogeological model provides nine sets for training (or validation, or testing). These statistics indicate that as the size of the training dataset increases, the results favor improved learning, with a decrease in MAE and RMSE, and an increase in R^2. A distinct pattern emerges as the number of epochs increases. The metrics undergo rapid changes initially, stabilizing and reaching a plateau after 300 epochs. This suggests that extending training beyond this point may not be necessary. This underscores the importance of diversity within the training dataset, which occurs when more models are provided. However, it is essential to note that this comes at a substantial cost in terms of training time, which is proportional to the number of epochs and the model size. For example, when the number of epochs doubles (and likewise for the model size), the training time also doubles. We, therefore, chose a configuration of 200 epochs and 400 models for training. This choice allowed us to keep training times at approximately 3 h while still achieving an adequate level of precision.

4.3. S-STBM results

Scenario 4 is employed for the inversion. The prior distributions for hydraulic conductivity, boundary conditions, specific storage, and pumping rate are assumed to be the same as those in Table 2.

One hundred realizations were calibrated using S-STBM with a parameter of $l = 100$ and 400 iterations for the inversion. Scatterplots of the reference values against the inversion results using U-Net and the flow simulator MRST are presented for two realizations, along with the mean calibrated heads and their standard deviation (refer to Fig. 8). The red dashed lines represent the 95% confidence interval corresponding to the Gaussian noise with $\sigma = 0.05$ added to the sampled reference data.

It can be observed that for both methods, the majority of points fall within the confidence interval for realizations 1 and 2, as well as for the average estimated heads. Slightly larger discrepancies were observed when using the flow simulator, but they were not sufficient to conclude that using the deep learning network would be more effective than an approach utilizing a flow simulator. These discrepancies are apparent through the slightly larger error bars in the MRST method. Consequently, both methods effectively condition the transient pressure heads for this problem.

The log-hydraulic conductivity fields obtained through inversion using S-STBM with both U-Net and MRST are presented in Fig. 9 for the same two realizations, alongside the mean log-hydraulic conductivity values and their corresponding variances. The results demonstrate that both S-STBM realizations accurately reproduce the spatial structure of the reference field. The mean log-conductivity fields reveal that S-STBM (with U-Net and MRST) are capable of identifying high conductivity zones (just below the well and in the lower-left corner) and lower conductivity areas (left and top sides). Note that U-Net displays slightly higher mean values in regions with high conductivity (indicated by red areas) and slightly lower means in regions with low conductivity (indicated by blue areas) compared to the inversion carried out with MRST. U-Net demonstrates similar variance around the pumping zone when compared to MRST. However, these results do not provide sufficient robustness to definitively assert the efficiency of U-Net over MRST. Consequently, the inversion results of S-STBM with U-Net are comparable to those obtained with S-STBM with MRST. Further investigations across various hydrogeological contexts will be essential to draw conclusions regarding the efficiency of U-Net compared to MRST.

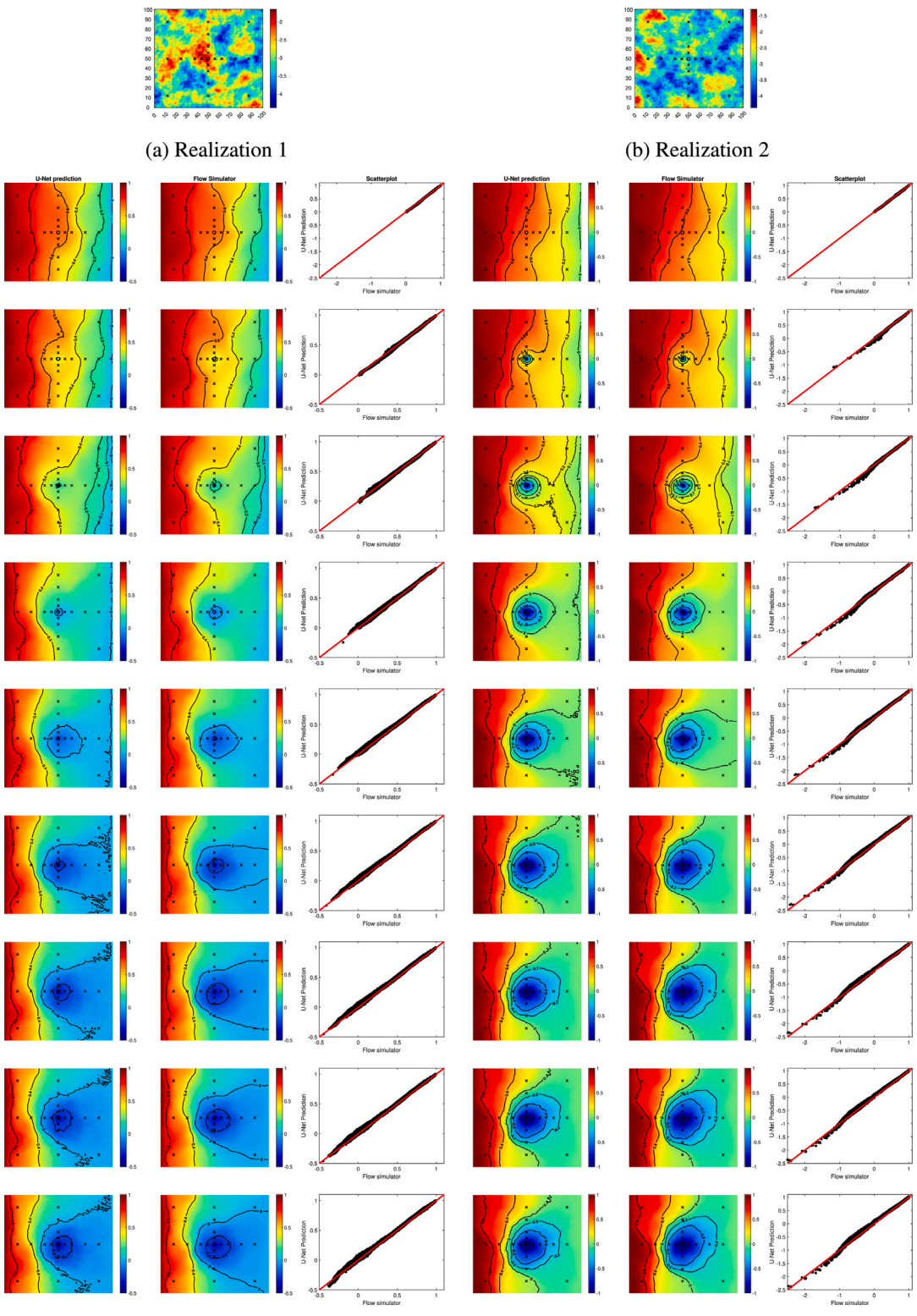


Fig. 6. Scenario (1) Validation images for visually evaluating the prediction performance of the trained U-Net. Two panels show the U-Net predictions (left), flow simulation (middle), and scatterplots (right) of actual head values against the U-Net predictions for a given realization 1, panel (a), and a given realization 2, panel (b). The corresponding times from top to bottom are: t = 0, 1, 5, 15, 30, 60, 120, 180, and 720 min.

Moreover, it is evident that the simulated conductivities in regions beyond the influence zone of the well, particularly near the boundaries where no piezometers are located, exhibit significant uncertainty. This underscores the difficulty of determining the actual hydraulic conductivity field using observed pressure head data, which is the sole available information. This situation generates an ill-posed problem

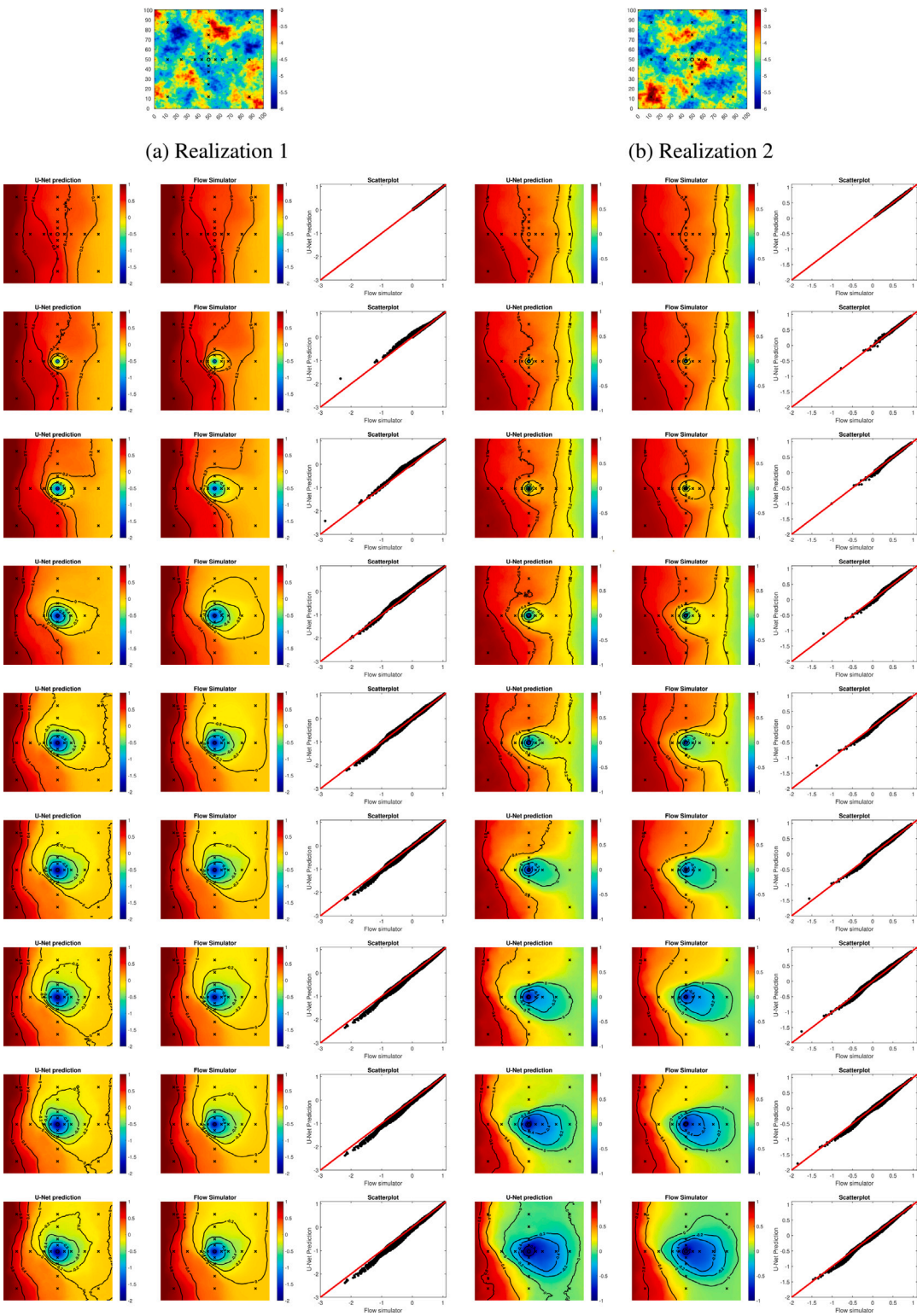


Fig. 7. Scenario (4) Validation images for visually evaluating the prediction performance of the trained U-Net. Two panels show the U-Net predictions (left), flow simulation (middle), and scatterplots (right) of actual head values against the U-Net predictions for a given realization 1, panel (a), and a given realization 2, panel (b). The corresponding times from top to bottom are: t = 0, 1, 5, 15, 30, 60, 120, 180, and 720 min.

within the hydrogeological system under study. One solution can be to couple the hydrogeological models with an available geophysical model (Neven and Renard, 2023).

The uncertainty associated with boundary conditions, specific storage, and pumping rate was also evaluated. In Table 5, the mean, standard deviation, and the 15–85th quantiles are presented. The performance of S-STBM with U-Net (up) is compared with S-STBM using MRST (bottom). The results indicate no significant difference between the U-Net approach and utilizing a flow simulator. Furthermore, the posterior distributions encompass the reference values, suggesting

Table 4
Scenario (4) Investigation of the performance metrics (MAE, RMSE, R^2) between observed and predicted U-Net heads across varying combinations of epoch counts and training dataset sizes. Additionally, the training times are presented.

Epoch	Training size	MAE (m)	RMSE (m)	R^2 (–)	Training time (h)
100	100	0.068	0.097	0.917	0.36
	200	0.062	0.086	0.930	0.64
	300	0.041	0.058	0.970	1.06
	400	0.031	0.046	0.983	1.35
	500	0.033	0.046	0.984	1.66
200	100	0.062	0.086	0.930	0.64
	200	0.056	0.076	0.951	1.31
	300	0.037	0.055	0.974	2.00
	400	0.033	0.047	0.983	2.95
	500	0.028	0.042	0.985	3.74
300	100	0.062	0.087	0.932	0.97
	200	0.049	0.069	0.957	2.00
	300	0.036	0.053	0.975	3.11
	400	0.030	0.045	0.982	3.98
	500	0.027	0.041	0.986	5.23
400	100	0.065	0.089	0.926	1.31
	200	0.053	0.074	0.955	3.52
	300	0.038	0.055	0.972	4.21
	400	0.032	0.048	0.982	5.34
	500	0.027	0.040	0.986	6.95
500	100	0.067	0.093	0.924	1.64
	200	0.050	0.070	0.957	3.41
	300	0.038	0.055	0.973	5.15
	400	0.033	0.049	0.980	6.97
	500	0.028	0.042	0.985	8.56

Table 5
Posterior distribution after solving the inversion problem using U-Net (top) and MRST (bottom). (Ref: Reference values, Mean: Mean value, Std: Standard deviation, Q15: 15th quantile, Q85: 85th quantile of one hundred realizations. *B.C.*: Boundary conditions, S_s: specific storage, Q: pumping rate.).

Model	Properties	Ref	Mean	Std	Q15	Q85
U-Net	*B.C.* left side (m)	−0.035	−0.038	0.023	−0.058	−0.017
	B.C. right side (m)	1.042	1.043	0.037	1.005	1.084
	$log10(S_s)$ (m^{-1})	−4.090	−4.110	0.102	−4.207	−4.009
	Q (L/min)	17	−17.187	2.010	−19.400	−14.85
MRST	*B.C.* left side (m)	−0.035	−0.044	0.033	−0.083	−0.007
	B.C. right side (m)	1.042	−1.050	0.040	1.019	1.0864
	$Log10(S_s)$ (m^{-1})	−4.090	−4.039	0.063	−4.113	−3.977
	Q (L/min)	17	−18.000	1.651	−19.458	−16.093

that both scenarios effectively retrieve parameter values and their uncertainties. The methodology demonstrates the capability to simultaneously determine a heterogeneous hydraulic conductivity field and single-value parameters.

An undeniable advantage of the S-STBM approach with U-Net, in comparison to the S-STBM approach with MRST, is the execution time of the forward model. With U-Net, the average execution time to obtain transient head fields is 0.304 s, in contrast to 3.509 s when using the MRST flow simulator—a difference of one order of magnitude. For this study, it takes approximately 3.57 h to obtain inversion results with U-Net and around 40 h hours when using MRST. Even taking into account the U-Net network training time of 3 h, S-STBM with U-Net remains six times faster in execution time while producing very similar results to the approach employing a flow simulator.

5. Discussion

A novel methodology has been implemented to solve inverse problems by combining S-STBM and U-Net for calibrating transient hydraulic heads. U-Net serves as an emulator, replacing the conventional physics-based model typically used in hydrogeological inverse problems. The model was trained on a dataset comprising 400 hydrogeological models and their corresponding flow simulations obtained through MRST. To expand the training set, an auto-regressive strategy was employed, incorporating hydraulic heads from the previous time step as input to predict heads at the current time step. Nine-time steps were utilized, resulting in the creation of 3600 (9 times 400) datasets for training. This approach enhances diversity and increases the size of the training dataset. Following training, U-Net was used to emulate the forward operator in the geostatistical S-STBM algorithm, enabling the quantification of uncertainties related to hydraulic conductivity.

The performance of U-Net was evaluated using a separate dataset containing 100 hydrogeological models (900 datasets) and corresponding flow simulations, as well as applied to a synthetic case study. The results demonstrate U-Net's ability to successfully establish a correspondence between given heterogeneous hydraulic conductivity field, boundary conditions, source term, and specific storage, enabling the prediction of corresponding transient flow simulations of a 12-hour pumping test. While the generated pressure maps exhibit slight noise compared to reference maps, there is a strong correlation between the predicted and actual transient hydraulic heads. These errors were small enough to have minimal impact on the inversion results. The hydraulic conductivity maps and their uncertainties obtained using U-Net closely match those obtained by S-STBM when used with a flow simulator. The study highlights U-Net's effectiveness in hydrogeological inversion of transient hydraulic heads, reducing computational demands compared to the MRST flow simulator.

A primary contribution of the proposed framework lies in its incorporation of boundary conditions, the source term, including intensity and activation time, and the specific storage as input data to preserve

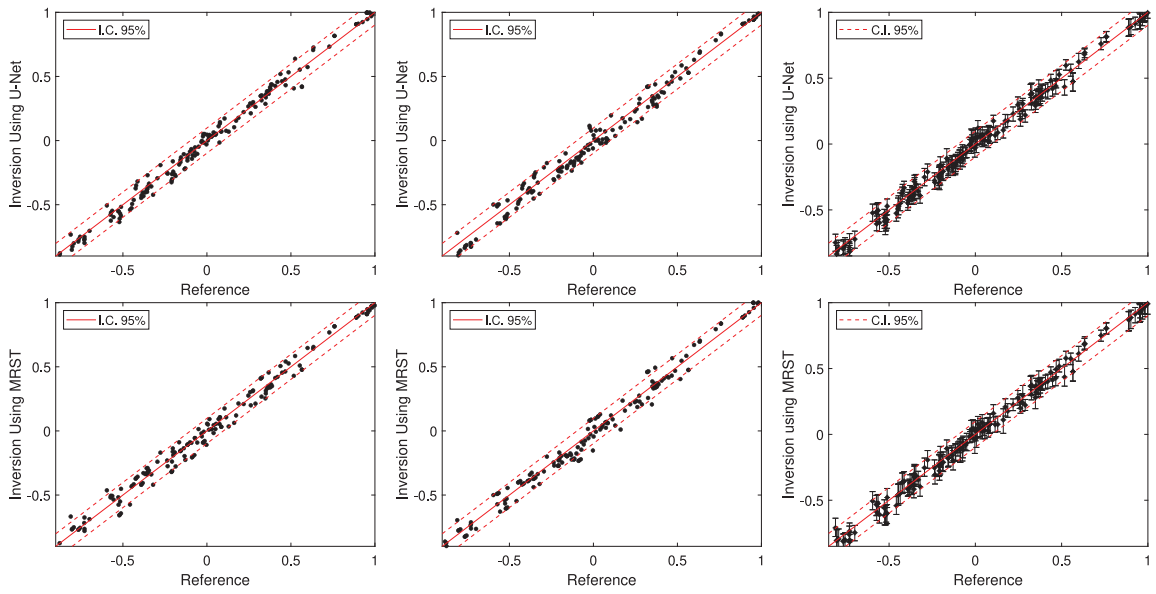


Fig. 8. Scatterplots depicting the relationship between actual head values and the inversions generated by the U-Net model (top) and the MTSR model (bottom) for two specific realizations (left side). On the right side, the scatterplot showcases the average head values across 100 realizations along with their associated uncertainties. (Red dashed lines: 95% confidence interval corresponding to Gaussian noise with a standard deviation of $\sigma = 0.05$).

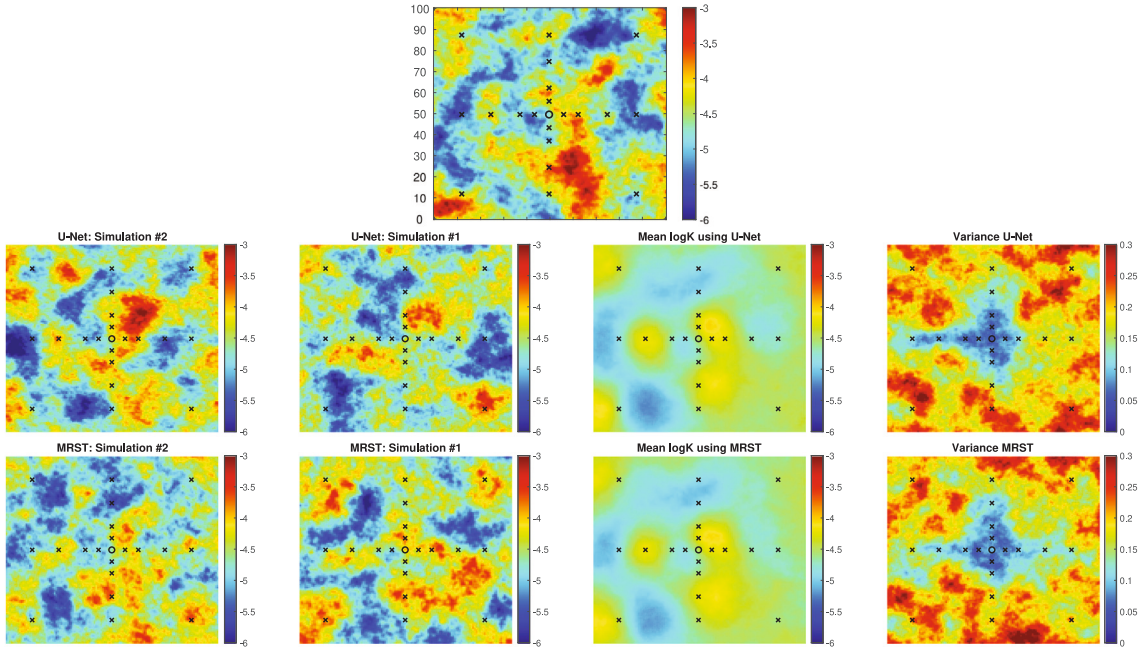


Fig. 9. The reference log-hydraulic conductivity field (Top). The hydraulic conductivity fields, derived from inversions performed by the U-Net model (middle) and the MRST model (bottom) for two distinct realizations, displayed on the left side. On the right side, the average hydraulic conductivity field computed across 100 realizations, accompanied by the corresponding variance field.

the model's physics. The U-Net network was effectively trained to respect Neumann boundary conditions, resulting in no flow at the northern and southern boundaries, as observed in Figs. 6 and 7. By including the source term in the network's input parameters, the temporal aspect of the problem is considered.

Surrogate modeling using U-Net extends beyond continuous fields. Categorical aquifers can be taken into account by using the approach developed by Lauzon and Marcotte (2023), which employs latent Gaussian fields to generate categorical variables, or the approach proposed by Dagasan et al. (2020), which employs multiple point simulations and a sampling technique to solve the inverse problem.

The simplicity of the proposed U-Net architecture has facilitated rapid training of the network. A pivotal factor contributing to this efficiency is the consistent maintenance of channel dimensions at 192, which corresponds to 32 feature maps for each input map ($192 = 32 \times 6$), in every downscale or upscale step within both the encoder and decoder. In contrast, the conventional approach of gradually doubling channel sizes up to 1024 channels, as commonly observed in U-Net architectures (Mo et al., 2019b; Taccari et al., 2022), would lead to a substantial increase in learnable variables from 10.0 million to hundreds of millions in our study. Such an expansion would impose considerable limitations on memory and time, thus motivating the

decision to maintain a fixed number of channels. This choice has demonstrated minimal impact on U-Net's prediction capability.

One strategy to enhance U-Net's predictive performance could involve incorporating an attention mechanism, as proposed by Taccari et al. (2022), aimed at improving the network's approximation accuracy and reducing model uncertainty by focusing on crucial areas. Nevertheless, it is important to acknowledge that this integration would inevitably extend the network's training time. This enhancement could potentially yield benefits for more complex problems compared to the one presented.

Our machine-learning network serves as a surrogate model to estimate responses across various geological scenarios, encompassing a range of hydraulic conductivity fields, boundary conditions, pumping wells, or specific storage conditions. However, our current approach does not account for scenarios in which hard data (i.e., known hydraulic conductivity at a specific cell) or soft data (e.g., geophysical survey or geophysical log data) are available. In such cases, we can tailor the training of the neural networks to address a specific geological problem by directly incorporating hard data and soft data into the models used for training. This streamlined approach involves creating training sets that accurately capture the behavior of both hard and soft data. Geostatistical algorithms can be employed to condition geological fields to the available hard and soft data (Azevedo and Soares, 2017; Lauzon and Marcotte, 2020a; Neven and Renard, 2023). This focused training can result in specialized networks designed for a particular site or project, potentially improving accuracy and reducing training time. The parameter space is constrained due to the conditioning effect of hard and soft data. Additionally, these specialized networks can be easily reused and adapted if additional data becomes available, particularly when coupled with geophysical data during the inversion process (Neven and Renard, 2023). Further research is required to fully understand and address these behaviors.

A situation not discussed pertains to the marginal distribution and variogram parameters of the hydraulic conductivity field, assumed to be known. However, in real-life applications, assessing these parameters is a complex task due to the potentially high uncertainty in the marginal distribution and variogram resulting from a lack of data. Handling this uncertainty in network training is straightforward; one can generate models with different marginal distributions and variogram structures to address the uncertainty. The mapping of hydraulic conductivity is thus performed in the original space, as opposed to mapping a zero-mean Gaussian random field with a fixed marginal distribution. However, this approach may require more models to obtain a satisfactory surrogate model, and further research needs to be conducted to verify the applications of such methodology. For the inverse problem, S-STBM can handle unknown variogram parameters by simulating an isotropic field with a unit range and transforming it using contraction, dilation, and rotation matrices with sets of variogram parameters. The only requirement is to fix a variogram model, in this case, the spherical model. The Gaussian anamorphosis is capable of handling any marginal distribution by transforming the original data space into a Gaussian distribution. To address uncertainty in the marginal distribution, it is necessary to use a parametrizable marginal distribution, allowing S-STBM to optimize its parameters.

6. Conclusion

This article introduces a U-Net convolutional encoder–decoder network, which efficiently calculates the response to the transient state of a groundwater system during a pumping test. The data-driven surrogate model is trained and tested using a limited number of models, where transient hydraulic heads throughout the domain need to be inferred. By training it to minimize the discrepancy with target images, the proposed U-Net model easily captures the highly nonlinear relationship between inputs (hydraulic conductivity fields, boundary conditions, source term location, and intensity, and specific storage) and the output (transient hydraulic head field). To expedite learning, an auto-regressive strategy is employed, where hydraulic heads from the previous time step serve as inputs to predict the current time step output. The performance of the proposed surrogate is demonstrated on a synthetic problem with a heterogeneous conductivity field with uncertain boundary conditions, pumping rate, and specific storage. The hydraulic conductivity maps and their uncertainties obtained using U-Net are found to be very similar to those obtained with a geostatistical inverse method, the S-STBM algorithm, when used with a flow simulator. The study has shown that U-Net can be successfully used for hydrogeological inversion of transient hydraulic heads with reduced computational demands compared to the use of a flow simulator.

CRediT authorship contribution statement

Dany Lauzon: Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of competing interest

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

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge the valuable guidance and practical expertise provided by Polytechnique Montréal.

Funding information

This research has received no external funding.

Code availability

The MATLAB Reservoir Simulation Toolbox is available at Lie (2019). The MATLAB computer codes for the U-Net architecture and the geostatistical inversion method, S-STBM, are available at https://github.com/Danlauz/UNET-SSTBM.

References

Asher, M.J., Croke, B.F.W., Jakeman, A.J., Peeters, L.J.M., 2015. A review of surrogate models and their application to groundwater modeling. Water Resour. Res. 51 (8), 5957–5973. http://dx.doi.org/10.1002/2015wr016967.

Ashworth, M., Elsheikh, A.H., Doster, F., 2022. Machine learning-based multiscale constitutive modelling: Development and application to dual-porosity mass transfer. Adv. Water Resour. 163, 104166. http://dx.doi.org/10.1016/j.advwatres.2022.104166.

Azevedo, L., Soares, A., 2017. Geostatistical Methods for Reservoir Geophysics. Springer International Publishing, http://dx.doi.org/10.1007/978-3-319-53201-1.

Bai, T., Tahmasebi, P., 2022. Characterization of groundwater contamination: A transformer-based deep learning model. Adv. Water Resour. 164, 104217. http://dx.doi.org/10.1016/j.advwatres.2022.104217.

Bárdossy, A., Hörning, S., 2015. Random mixing: An approach to inverse modeling for groundwater flow and transport problems. Transp. Porous Media 114 (2), 241–259. http://dx.doi.org/10.1007/s11242-015-0608-4.

Benoit, N., Marcotte, D., Boucher, A., D'Or, D., Bajc, A., Rezaee, H., 2017. Directional hydrostratigraphic units simulation using MCP algorithm. Stoch. Environ. Res. Risk Assess. 32 (5), 1435–1455. http://dx.doi.org/10.1007/s00477-017-1506-9.

Benoit, N., Marcotte, D., Molson, J., 2020. Stochastic correlated hydraulic conductivity tensor calibration using gradual deformation. J. Hydrol. 594, 125880. http://dx.doi.org/10.1016/j.jhydrol.2020.125880.

Carrera, J., Alcolea, A., Medina, A., Hidalgo, J., Slooten, L.J., 2005. Inverse problem in hydrogeology. Hydrogeol. J. 13 (1), 206–222. http://dx.doi.org/10.1007/s10040-004-0404-7.

Chilès, J.P., Delfiner, P., 1997. Discrete exact simulation by the Fourier method. In: Baafi, E., Schofield, N. (Eds.), In: Geostatistics Wollongong, Vol. 96, Kluwer Academic, Dordrecht, pp. 258–269.

Dagasan, Y., Juda, P., Renard, P., 2020. Using generative adversarial networks as a fast forward operator for hydrogeological inverse problems. Groundwater 58 (6), 938–950. http://dx.doi.org/10.1111/gwat.13005.

Deutsch, C., 1992. Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data (Ph.D. thesis). Stanford University.

Dietrich, C.R., Newsam, G.N., 1993. A fast and exact method for multidimensional Gaussian stochastic simulations. Water Resour. Res. 29 (8), 2861–2869. http://dx.doi.org/10.1029/93wr01070.

Emery, X., Arroyo, D., Porcu, E., 2016. An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields. Stoch. Environ. Res. Risk Assess. 30 (7), 1863–1873. http://dx.doi.org/10.1007/s00477-015-1151-0.

Godoy, V.A., Napa-García, G.F., Gómez-Hernández, J.J., 2022. Ensemble random forest filter: An alternative to the ensemble Kalman filter for inverse modeling. J. Hydrol. 615, 128642. http://dx.doi.org/10.1016/j.jhydrol.2022.128642.

He, Q., Barajas-Solano, D., Tartakovsky, G., Tartakovsky, A.M., 2020. Physics-informed neural networks for multiphysics data assimilation with application to subsurface transport. Adv. Water Resour. 141, 103610. http://dx.doi.org/10.1016/j.advwatres.2020.103610.

Hu, 2000. Gradual deformation and iterative calibration of Gaussian-related stochastic models. Math. Geol. 32 (1), 87–108.

Jäggli, C., Straubhaar, J., Renard, P., 2017. Posterior population expansion for solving inverse problems. Water Resour. Res. 53 (4), 2902–2916. http://dx.doi.org/10.1002/2016wr019550.

Jiang, Z., Tahmasebi, P., Mao, Z., 2021. Deep residual U-net convolution neural networks with autoregressive strategy for fluid flow predictions in large-scale geosystems. Adv. Water Resour. 150, 103878. http://dx.doi.org/10.1016/j.advwatres.2021.103878.

Khambhammettu, P., Renard, P., Doherty, J., 2020. The traveling pilot point method. a novel approach to parameterize the inverse problem for categorical fields. Adv. Water Resour. 138, 103556. http://dx.doi.org/10.1016/j.advwatres.2020.103556.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. http://dx.doi.org/10.48550/ARXIV.1412.6980.

Laloy, E., Hérault, R., Lee, J., Jacques, D., Linde, N., 2017. Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network. Adv. Water Resour. 110, 387–405. http://dx.doi.org/10.1016/j.advwatres.2017.09.029.

Lantuéjoul, C., 2002. Geostatistical Simulation. Springer Berlin Heidelberg, http://dx.doi.org/10.1007/978-3-662-04808-5.

Lauzon, D., Marcotte, D., 2020a. Calibration of random fields by a sequential spectral turning bands method. Comput. Geosci. 135, 104390. http://dx.doi.org/10.1016/j.cageo.2019.104390.

Lauzon, D., Marcotte, D., 2020b. The sequential spectral turning band simulator as an alternative to Gibbs sampler in large truncated– or pluri– Gaussian simulations. Stoch. Environ. Res. Risk Assess. 34 (11), 1939–1951. http://dx.doi.org/10.1007/s00477-020-01850-9.

Lauzon, D., Marcotte, D., 2022. Statistical comparison of variogram–based inversion methods for conditioning to indirect data. Comput. Geosci. 105032. http://dx.doi.org/10.1016/j.cageo.2022.105032.

Lauzon, D., Marcotte, D., 2023. Joint hydrofacies-hydraulic conductivity modeling based on a constructive spectral algorithm constrained by transient head data. Hydrogeol. J. http://dx.doi.org/10.1007/s10040-023-02638-1.

Le Ravalec, M., Noetinger, B., Hu, L.Y., 2000. The FFT moving average generator : An efficient numerical method for generating and conditioning Gaussian simulations. Math. Geol. 32 (6), 701–722.

Li, J., Lu, W., Luo, J., 2021. Groundwater contamination sources identification based on the long-short term memory network. J. Hydrol. 601, 126670. http://dx.doi.org/10.1016/j.jhydrol.2021.126670.

Lie, K.-A., 2019. An Introduction to Reservoir Simulation Using MATLAB/GNU Octave. Cambridge University Press, http://dx.doi.org/10.1017/9781108591416, Title : MRST Transforming research. URL: https://www.sintef.no/projectweb/mrst/, (Accessed 12 December 2022).

Luenberger, D., Ye, Y., 2008. Linear and Nonlinear Programming. Springer US, URL: https://www.ebook.de/de/product/19292757/david_g_luenberger_yinyu_ye_linear_and_nonlinear_programming.html.

Mariethoz, G., Gómez-Hernández, J.J., 2021. Editorial: Machine learning for water resources. Front. Artif. Intell. 4, http://dx.doi.org/10.3389/frai.2021.699862.

Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-Net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision. 3DV, IEEE, pp. 1–11. http://dx.doi.org/10.1109/3dv.2016.79.

Mo, S., Zabaras, N., Shi, X., Wu, J., 2019a. Deep autoregressive neural networks for high–dimensional inverse problems in groundwater contaminant source identification. Water Resour. Res. 55 (5), 3856–3881. http://dx.doi.org/10.1029/2018wr024638.

Mo, S., Zhu, Y., Zabaras, N., Shi, X., Wu, J., 2019b. Deep convolutional encoder-decoder networks for uncertainty quantification of dynamic multiphase flow in heterogeneous media. Water Resour. Res. 55 (1), 703–728. http://dx.doi.org/10.1029/2018wr023528.

Mosegaard, K., Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. J. Geophys. Res.: Solid Earth 100 (B7), 12431–12447. http://dx.doi.org/10.1029/94jb03097.

Neven, A., Renard, P., 2023. A novel methodology for the stochastic integration of geophysical and hydrogeological data in geologically consistent models. Water Resour. Res. 59 (7), http://dx.doi.org/10.1029/2023wr034992.

Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., Glocker, B., Rueckert, D., 2018. Attention U-net: Learning where to look for the pancreas. http://dx.doi.org/10.48550/ARXIV.1804.03999.

Pasquier, P., Marcotte, D., 2006. Steady- and transient-state inversion in hydrogeology by successive flux estimation. Adv. Water Resour. 29 (12), 1934–1952. http://dx.doi.org/10.1016/j.advwatres.2006.02.001.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: Lecture Notes in Computer Science. Springer International Publishing, pp. 234–241. http://dx.doi.org/10.1007/978-3-319-24574-4_28.

Secci, D., A. Godoy, V., Gómez-Hernández, J.J., 2024. Physics-informed neural networks for solving transient unconfined groundwater flow. Comput. Geosci. 182, 105494. http://dx.doi.org/10.1016/j.cageo.2023.105494.

Shinozuka, M., 1971. Simulation of multivariate and multidimensional random processes. J. Acoust. Soc. Am. 49 (1B), 357–368. http://dx.doi.org/10.1121/1.1912338.

Shinozuka, M., Deodatis, G., 1996. Simulation of multi-dimensional Gaussian stochastic fields by spectral representation. Appl. Mech. Rev. 49 (1), 29–53. http://dx.doi.org/10.1115/1.3101883.

Shinozuka, M., Jan, C.M., 1972. Digital simulation of random processes and its applications. J. Sound Vib. 25, 111–128.

Siddique, N., Paheding, S., Elkin, C.P., Devabhaktuni, V., 2021. U-Net and its variants for medical image segmentation: A review of theory and applications. IEEE Access 9, 82031–82057. http://dx.doi.org/10.1109/access.2021.3086020.

Taccari, M.L., Nuttall, J., Chen, X., Wang, H., Minnema, B., Jimack, P.K., 2022. Attention U–Net as a surrogate model for groundwater prediction. Adv. Water Resour. 163, 104169. http://dx.doi.org/10.1016/j.advwatres.2022.104169.

Tang, M., Liu, Y., Durlofsky, L.J., 2020. A deep-learning-based surrogate model for data assimilation in dynamic subsurface flow problems. J. Comput. Phys. 413, 109456. http://dx.doi.org/10.1016/j.jcp.2020.109456.

Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics, http://dx.doi.org/10.1137/1.9780898717921.

Tripathy, R.K., Bilionis, I., 2018. Deep UQ: Learning deep neural network surrogate models for high dimensional uncertainty quantification. J. Comput. Phys. 375, 565–588. http://dx.doi.org/10.1016/j.jcp.2018.08.036.

Wang, N., Chang, H., Zhang, D., 2021. Deep–learning–based inverse modeling approaches: A subsurface flow example. J. Geophys. Res.: Solid Earth 126 (2), http://dx.doi.org/10.1029/2020jb020549.

Wang, N., Chang, H., Zhang, D., 2022. Surrogate and inverse modeling for two–phase flow in porous media via theory–guided convolutional neural network. J. Comput. Phys. 466, 111419. http://dx.doi.org/10.1016/j.jcp.2022.111419.

Wu, J., Zhou, S., Zuo, S., Chen, Y., Sun, W., Luo, J., Duan, J., Wang, H., Wang, D., 2021. U-Net combined with multi-scale attention mechanism for liver segmentation in CT images. BMC Med. Inform. Decis. Mak. 21 (1), http://dx.doi.org/10.1186/s12911-021-01649-w.

Zhang, X., Zhu, Y., Wang, J., Ju, L., Qian, Y., Ye, M., Yang, J., 2022. GW-PINN: A deep learning algorithm for solving groundwater flow equations. Adv. Water Resour. 165, 104243. http://dx.doi.org/10.1016/j.advwatres.2022.104243.

Zhu, Y., Zabaras, N., 2018. Bayesian deep convolutional encoder–decoder networks for surrogate modeling and uncertainty quantification. J. Comput. Phys. 366, 415–447. http://dx.doi.org/10.1016/j.jcp.2018.04.018.