arXiv:2401.04823v2 [math.AP] 23 Apr 2024

Deep learning surrogate for predicting hydraulic conductivity tensors from DFM models

Martin Špetlík[1*], Jan Březina[1] and Eric Laloy[2]

[1*]Institute of New Technologies and Applied Informatics, Faculty of Mechatronics, Informatics and Interdisciplinary Studies, Technical University of Liberec, Studentská 1402/2, Liberec, 461 17, Czech Republic.

[2]Engineered and Geosystems Analysis Unit, Sustainable Waste Management and Decommissioning, Belgian Nuclear Research Center (SCK CEN), Boeretang 200, Mol, 2400, Belgium.

*Corresponding author(s). E-mail(s): martin.spetlik@tul.cz;
Contributing authors: jan.brezina@tul.cz; eric.laloy@sckcen.be;

Abstract

Simulating water flow in fractured crystalline rock requires tackling its stochastic nature. We aim to utilize the multilevel Monte Carlo method for cost-effective estimation of simulation statistics. This multiscale approach entails upscaling of fracture hydraulic conductivity by homogenization. In this work, we replace 2D numerical homogenization based on the discrete fracture-matrix (DFM) approach with a surrogate model to expedite computations. We employ a deep convolutional neural network (CNN) connected to a deep feed-forward neural network as the surrogate. The equivalent hydraulic conductivity tensor $\boldsymbol{K^{eq}}$ is predicted based on the input tensorial spatial random fields (SRFs) of hydraulic conductivities, along with the cross section and hydraulic conductivity of fractures. Three independent surrogates with the same architecture are trained, each with a different ratio of fracture to matrix hydraulic conductivity $\boldsymbol{K_f/K_m}$. As the ratio $\boldsymbol{K_f/K_m}$ increases, the distribution of $\boldsymbol{K^{eq}}$ becomes more complex, leading to a decline in the prediction accuracy of the surrogates. The prediction accuracy improves as the fracture density decreases, regardless of the $\boldsymbol{K_f/K_m}$. We also investigate prediction accuracy for different correlation lengths of input SRFs. The observed speedup gained by surrogates varies from $\mathbf{4}\times$ to $\mathbf{28}\times$ depending on the number of homogenization blocks. Upscaling by numerical homogenization and surrogate modeling is compared on two macroscale problems. For the first one, the accuracy of outcomes is directly correlated with the precision of $\boldsymbol{K^{eq}}$ predictions. For the latter one, we observe only a mild impact of the upscaling method on the accuracy of the results.

Keywords: Deep learning surrogate, 2D DFM models, Numerical homogenization, Equivalent hydraulic conductivity

1 Introduction

Modeling groundwater processes is integral to researching deep geological repositories of radioactive waste. The repository planned in the Czech Republic will be located in crystalline rock, where a small intergranular pore space causes groundwater to flow predominantly through a fracture network [1]. Given the inherent uncertainties in the modeled processes, our aim is to utilize the

multilevel Monte Carlo method (MLMC) [2] to obtain statistics at a reduced computational cost compared to the plain Monte Carlo method. Applying MLMC directly to models with discrete fractures, such as discrete fracture networks (DFN), poses challenges. Coarser-level models in MLMC cannot capture fractures below their resolution, and DFNs themselves struggle to resolve fractures smaller than about 1/1000th of the domain size due to an exponential growth in fracture count. The impact of the small fractures, however, should not be ignored. To address this limitation, we adopt a discrete fracture-matrix (DFM) approach that integrates both continuum and discrete fractures. This allows us to homogenize the impact of small fractures into the matrix properties when transitioning from a finer to a coarser scale DFM model. A surrogate based on deep learning techniques can accelerate a costly numerical homogenization procedure. This study investigates deep learning-based surrogates for the numerical homogenization of an equivalent hydraulic conductivity tensor for 2D DFM models.

Homogenization techniques typically utilize a representative elementary volume (REV) to calculate the effective hydraulic conductivity of fractured media [3]. However, REV is not applicable for fracture networks in rocks due to their fractal structure [4]. The homogenized hydraulic conductivity becomes a heterogeneous random field with an increased lower bound for correlation length. In such cases, the terminology shifts to equivalent properties instead of effective properties (see [5, 6]). For our work, we utilize numerical homogenization in a block-averaged sense, where the correlation length of the resulting field varies depending on the block size (see [7]). While some authors differentiate between homogenization and upscaling, e.g., [8], we use them interchangeably in this work.

Our objective is to apply numerical homogenization to calculate an equivalent hydraulic conductivity tensor. The determination of the equivalent hydraulic conductivity of fractured porous media has been reviewed in [5, 6, 9]. Additionally, numerical homogenization in the context of DFM models has also been studied. Bogdanov et al. [10] examine a cubic rock matrix containing single-size two-dimensional polygonal fractures. The rock matrix is assumed to be isotropic and homogeneous. As an extension of the work by Koudina et al. [11], they employ two distinct permeabilities: one for the fractures and another for the rock matrix. A tetrahedral mesh with fractures is utilized to model the Darcy flow separately for the porous matrix and the fracture network. Bogdanov and colleagues also investigate the impact of fracture density and matrix hydraulic conductivity on the upscaled hydraulic conductivity. Their subsequent work [12] extends the previous work by considering power-law distributed fracture sizes. Lang et al. [13] determine an equivalent hydraulic conductivity tensor using a pressure gradient and a flux averaging technique for both 2D and 3D DFM models on unstructured grids. Other relevant studies include, for example, [14, 15].

Deep learning (DL) [16] is an extensive part of machine learning that utilizes deep architectures of artificial neural networks. In our case, deep neural networks (DNNs) are employed to solve regression problems using data-driven supervised learning [16]. An unknown mapping of input data to output data is learned from collected sample pairs of inputs and corresponding outputs of a DFM model. The input comprises a spatial random field (SRF) on an unstructured mesh, along with a discrete fracture network with random properties. The output corresponds to an equivalent hydraulic conductivity tensor. Deep learning methods have recently proven successful in determining the properties of porous media. Specifically, supervised convolutional neural networks (CNNs) have gained popularity for this purpose. Studies such as [17], [18] or [19] adopt 2D CNNs to predict permeability or porosity based on input images of a porous media. Many researchers [20–24] focus on more realistic 3D models, where 3D CNNs are employed. Hong and Liu [21] train four 3D CNNs for four different target permeabilities. They consider 3D binary images of the pore geometry as input and four permeabilities (three axial directions and mean permeability) as the corresponding output. Meng et al. [25] introduced a CNN with a transformer neural network to predict permeability. To predict equivalent permeability tensors, various approaches have been employed, including deep feed-forward neural networks [26], [27], graph convolutional neural networks [28], and generative adversarial networks [29]. Although He et al. [30] and Andrianov [31] use CNNs to upscale DFM models, they consider constant hydraulic conductivity for both the matrix and fractures.

This approach is inadequate for our needs, as we intend to incorporate our surrogate into stochastic calculations using MLMC.

Our work represents a pioneering effort in employing surrogates to predict equivalent hydraulic conductivity tensors from 2D DFM models, where both matrix and fracture elements exhibit non-constant hydraulic conductivity. Our study on 2D surrogates serves as a proof of concept before dealing with the more applicable but more complex 3D homogenization case. Our contribution involves the utilization of tensorial spatial random fields (SRFs) for matrix conductivities and non-constant hydraulic conductivity of fractures, which depends on fracture size and aperture. We investigate the impact of the matrix-to-fracture hydraulic conductivity ratio on the distribution of the equivalent hydraulic conductivity tensor and the associated predictive capabilities of the trained surrogates. The accuracy of the trained surrogates is further analyzed using unseen DFM model data with varying SRF correlation lengths or number of fractures. Finally, we compare the trained surrogates to numerical homogenization using the resulting macroscale hydraulic conductivity field within two macroscale problems: groundwater flow through a given surface (Aquifer problem) and upscaling the hydraulic conductivity tensor (Anisotropy problem).

The article is structured as follows. Section 2 explains multiscale DFM models as a conjunction of a discrete fracture network (DFN) and an equivalent continuous media (ECM) approach. The numerical homogenization of the equivalent hydraulic conductivity tensor is described, including two investigated macroscale problems. The numerical homogenization is put into the context of multiscale models. Section 3 provides the surrogate architecture and metrics for assessing surrogate accuracy. Section 4 describes the dataset configurations and preprocessing adopted. Section 5 presents and discusses the results, including the remaining challenges. Finally, Section 6 provides the conclusions of the article.

2 Multiscale models of fractured media

The primary aim of this section is to introduce homogenization problems, which will later be replaced by surrogates. The second aim is to contextualize these homogenization problems within a multiscale description of fractured media and the multilevel Monte Carlo method. We start with a simple DFN stochastic model of the distribution of discrete fractures. Then, we proceed to the ECM description of the rock mass by a random permeability tensor field. We then combine these descriptions using a DFM approach. Refer to [32] for various descriptions of fractured rock. We continue with homogenization and two macroscale test problems. Finally, we describe the motivation behind homogenization using fast surrogates to enable the multilevel Monte Carlo method for Darcy flow in fractured rocks.

2.1 Discreate fracture network

As Darcy flow in crystalline rock is primarily attributed to fractures, discrete fracture network (DFN) models are commonly employed: [33], [34], [35]. The characteristics of fractures in DFNs can include (1) length, (2) aperture, (3) orientation, or (4) density (as well as connectivity), all of which are regarded as random variables.

1. Based on evidence from nature, it is commonly assumed that fracture lengths (f_s) follow a power-law distribution (see [4]):

$$f_s \sim C r^{-\alpha}, \qquad C = \frac{1-\alpha}{\overline{r}^{1-\alpha} - \underline{r}^{1-\alpha}}, \tag{1}$$

   where α represents the characteristic exponent, while $\underline{r}$ and $\overline{r}$ denote the minimum and maximum lengths of a fracture, respectively.
2. The aperture of a fracture plays a crucial role in determining its permeability. According to Bonnet et al. [4], various distributions such as log-normal, power-law, normal, or Lévy-stable have been utilized to describe the aperture distribution. However, for simplification purposes, some studies (e.g., [13, 36]) adopt a constant aperture for all fractures. We assume a linear relation between fracture aperture and size: $\delta = a f_s$, with $a = 1 \times 10^{-4}$.

3. The distribution of orientations is frequently modeled using von Mises or Fisher distributions (for further details, refer to [37]). For the sake of simplicity, we consider an isotropic distribution of orientations.
4. Dimensionless density, as described in [37], serves to quantify the degree of connectivity of fractures. It determines the number of fracture centers (i.e., the total number of fractures) within a given domain size. In the 2D case, the dimensionless density is defined as follows:

$$\rho'_{2D} = \rho_{2D} A_{ex}, \tag{2}$$

where ρ_{2D} is the number of fractures per unit area, and A_{ex} denotes the excluded area, i.e., an area around a fracture where the center of another fracture cannot be located to ensure the two fractures do not intersect, for details see [38, ch. 3]. The concept of dimensionless density can be extended naturally to the 3D case with 2D fractures.

For simplicity, this article assumes uncorrelated fractures with positions determined by a Poisson process and independent orientations. See [38, p. 31] and [39] for more general models. Additionally, we consider the cubic law for fracture hydraulic conductivity:

$$K_f = \frac{g\rho_w}{12\mu}\delta^2, \tag{3}$$

where g is the gravitational acceleration, ρ_w denotes the density of water, and μ stands for the dynamic viscosity of water.

2.2 Spatially correlated random field on the matrix

The DFN approach becomes computationally prohibitive when dealing with a large number of fine-scale fractures. The equivalent continuous media (ECM) approach, as detailed in [40] and [41], offers a solution by representing rock hydraulic properties with a random permeability tensor field $\boldsymbol{K}(x)$. This random field can be characterized either implicitly, through the homogenization of a DFN realization (refer to Section 2.3), or explicitly, by describing its point distribution and correlation structure. In particular, we consider the permeability tensor field in the form:

$$\boldsymbol{K} = Q^T \Lambda Q, \quad Q = \begin{pmatrix} Y_x & -Y_y \\ Y_y & Y_x \end{pmatrix}, \quad \Lambda = \begin{pmatrix} k_x & 0 \\ 0 & k_y \end{pmatrix}.$$

We assume the isotropic rotation matrix Q, which is consistent with the isotropic distribution of fractures. In order to prescribe the rotation with a correlation structure, we use unit vector field $\mathbf{Y} = \mathbf{X}/|\mathbf{X}|$, where $\mathbf{X} = (X_x, X_y)$ has components that are independent Gaussian fields with zero mean, unit variance, and a specified correlation length λ. The principal permeabilities are given by the Gaussian vector $(k_x, k_y) = \exp\left(\boldsymbol{\mu} + \sqrt{\boldsymbol{\Sigma}}\,\tilde{\mathbf{k}}\right)$, where $\boldsymbol{\mu}$ is the mean vector, $\boldsymbol{\Sigma}$ is the covariance matrix, and $\tilde{\mathbf{k}}$ is a Gaussian field. The components of $\tilde{\mathbf{k}}$ are independent, follow a standard $N(0,1)$ distribution, and share a correlation length λ. Consequently, the tensor field is defined by four independent scalar fields: X_x, X_y, $\tilde{k}_x$, and $\tilde{k}_y$, each following a $N(0,1)$ distribution with a Gaussian correlation of length λ. For sampling these fields, the GSTools library by [42] is utilized.

2.3 Discrete fracture-matrix model

The discrete fracture-matrix (DFM) approach, as outlined in [43], [44], and [45], integrates aspects of both DFN and ECM through a mixed mesh composed of both 2D and 1D (or 3D and 2D) finite elements. The computational domain is decomposed into the set of fractures and the matrix domain: $\Omega = \Omega_f \sqcup \Omega_m$. In Ω_m, we consider Darcy flow equation:

$$\operatorname{div} \delta_m \mathbf{u}_m = 0, \quad \mathbf{u}_m = -\boldsymbol{K}_m \nabla h_m, \tag{4}$$

where $\mathbf{u}_m(\mathbf{x})$ denotes the Darcy velocity $[m/s]$ and $\boldsymbol{K}_m(\mathbf{x})$ is the permeability tensor $[m^2]$. To emphasize the true scaling of the quantities with respect to the characteristic length L, we introduce a non-physical principal unknown $h_m = p_m/\mu$ $[1/s]$, which is associated with the pressure p_m $[Pa]$ and the dynamic viscosity μ $[Pa \cdot s]$. Although the usage of permeability is good to justify the scaling, for the sake of practicality, we will henceforth use hydraulic conductivity out of this section, retaining the notation K. For the sake of simplicity of the homogenization formulas and consistency with the equation on Ω_f, we set $\delta_m = 1$.

The corresponding equation on Ω_f is given by:

$$\operatorname{div} \delta_f \mathbf{u}_f = q^+ + q^-, \quad \mathbf{u}_f = -k_f \nabla h_f, \quad (5)$$

where $\delta_f(\mathbf{x}) = ar_i$ $[m]$ is the aperture of fracture i and $k_f(\mathbf{x}) = \delta^2/12$ denotes the scalar fracture permeability. The sources $q^\pm$ denote outflows from the two aligned boundaries of Ω_m with normal vectors $\mathbf{n}^\pm$, coupling the fracture and matrix domain through Robin-like boundary conditions:

$$-\boldsymbol{K}_m \nabla h_m \cdot \mathbf{n}^\pm = q^\pm := \frac{k_f}{2\delta}(h_m^\pm - h_f) \quad \text{on } \Omega_f. \quad (6)$$

Our proprietary multiphysical software, Flow123d ([46]), is employed to solve DFM Darcy flow problems using mixed finite elements.

We now introduce two benchmark problems designed to assess the performance of the studied homogenization surrogates. Importantly, the second benchmark problem serves a dual role, also acting as the microscale homogenization problem to be approximated by the surrogates. The multiscale nature of this approach is further elaborated in the subsequent section.

2.3.1 Aquifer problem

We examine fluid flow within a square domain $\Omega = (0, L)^2$, where no-flow boundary conditions are applied on the top and bottom sides of the domain. We impose Dirichlet boundary conditions on the vertical sides, setting the hydraulic head $h = H$ at $x = 0$ and $h = 0$ at $x = L$. Let Y $[m^2 s^{-1}]$ denote the total horizontal flux through the domain. The equivalent horizontal permeability k_x can then be calculated as:

$$k_x = \frac{Y}{H}. \quad (7)$$

Clearly, Y is linear with respect to the prescribed pressure gradient H/L, so k_x is independent of H. Next, using unit analysis, we obtain the following scale equivariance: $k_x[L, L^2\boldsymbol{K}_m, L^2 k_f] = L^2 k_x[1, \boldsymbol{K}_m, k_f]$.

2.3.2 Anisotropy problem

To obtain the full permeability tensor, we solve two problems on $\Omega = (0, L)^2$, applying the boundary pressure head $P^1 = x$ and $P^2 = y$, respectively. Mixed formulation with RT_0 finite elements provides velocity vectors $\mathbf{u}(e, P)$ constant on the elements. Using element permeability tensors, we can approximate the pressure gradients as $\nabla h(e, P) = -\boldsymbol{K}_e^{-1}\mathbf{u}(e, P)$. For a vector quantity $\mathbf{v}(e, P)$ on elements, we introduce a weighted average:

$$\langle \mathbf{v} \rangle^i = \frac{\sum_{e \in \mathcal{T}} |e| \delta_e \mathbf{v}(e, P^i)}{\sum_{e \in \mathcal{T}} |e| \delta_e}. \quad (8)$$

We fit the homogenized symmetric tensor $\boldsymbol{K}^{eq}$ by solving the least squares problem:

$$- \begin{bmatrix} \langle \nabla h \rangle_x^1 & \langle \nabla h \rangle_y^1 & 0 \\ 0 & \langle \nabla h \rangle_x^1 & \langle \nabla h \rangle_y^1 \\ \langle \nabla h \rangle_x^2 & \langle \nabla h \rangle_y^2 & 0 \\ 0 & \langle \nabla h \rangle_x^2 & \langle \nabla h \rangle_y^2 \end{bmatrix} \begin{bmatrix} k_{xx}^{eq} \\ k_{xy}^{eq} \\ k_{yy}^{eq} \end{bmatrix} = \begin{bmatrix} \langle u \rangle_x^1 \\ \langle u \rangle_y^1 \\ \langle u \rangle_x^2 \\ \langle u \rangle_y^2 \end{bmatrix} \quad (9)$$

Note that $\langle \nabla h \rangle_x^1 \approx \langle \nabla h \rangle_y^2 \approx 1$ and $\langle \nabla h \rangle_y^1 \approx \langle \nabla h \rangle_x^2 \approx 0$, due to prescribed boundary conditions P^1 and P^2. Similarly to the previous case, the homogenization problem exhibits scale equivariance:

$$\boldsymbol{K}^{eq}[L, L^2\boldsymbol{K}_m, L^2 k_f] = L^2 \boldsymbol{K}^{eq}[1, \boldsymbol{K}_m, k_f] \quad (10)$$

2.4 Numerical homogenization in the context of multiscale models

For the benchmark problems 2.3.1 and 2.3.2, a fine mesh discretization can be employed to explicitly resolve all fractures within the size range $(\underline{r}, \overline{r}) = (h, L)$. A locally refined mesh with an element size upper bound $h \ll L$ is necessary for discretization compatible with fractures, as depicted in Fig. 1, left. We denote the problem result by $Q_h(\mathcal{F}_{h,L}, K_h)$, where $\mathcal{F}_{h,L}$ represents a specific realization of the fracture network for this size range, and $\boldsymbol{K}_h$ denotes a realization of the permeability random field effect of fractures up to size h. Considering the high computational cost of fine discretizations, a two-scale approach is employed as an alternative. This consists of a macro problem $Q_H(\mathcal{F}_{H,L}, \boldsymbol{K}_H)$ using a coarse mesh with maximal element size H, where $H > h$, and the coarse permeability tensor field $\boldsymbol{K}_H$ determined in terms of the micro problem 2.3.2.

In order to be compatible with convolutional neural networks, the numerical homogenization

problems are solved on a grid of homogenization blocks, similar to [47]. Four homogenization blocks are illustrated in Fig. 1 with center points $\mathbf{x}$, size $l \times l$, $l = 1.5H$, and overlap $l/2$. In order to evaluate boundary blocks, both fracture networks and the permeability field $\boldsymbol{K}_h$ are generated on a slightly larger domain of side $L + 2l$. For each block, we determine the intersection of the fractures $\mathcal{F}_{h,H}$ with the block, construct a compatible mesh with a maximal element size of h, and interpolate the field $\boldsymbol{K}_h$ to it. Then we calculate, the block's equivalent permeability $\boldsymbol{K}_H(\mathbf{x}) = P_h(\mathbf{x}, \mathcal{F}_{h,H}, \boldsymbol{K}_h)$. Finally, linear interpolation is used to transfer raster $\boldsymbol{K}_H$ values to the coarse grid cells, as exemplified in Fig. 1, right.

The two-scale procedure described will serve as a testbed for simulating the application of developed neural network surrogates, specifically designed to replace the homogenization problem P_h. While the acceleration of the two-scale model offered by homogenization surrogates is valuable on its own, the primary motivation lies in their application within the multilevel Monte Carlo (MLMC) method, as introduced by [2]. The MLMC estimate of the expectation $\mathbb{E}[Q]$ is based on a telescoping sum:

$$\langle Q \rangle_{N_1,\dots,N_I} = \sum_{i=1}^{I} \langle Q^{f,i} - Q^{c,i} \rangle_{N_i}, \qquad (11)$$

where N_i denotes the number of samples at resolution level i and $\langle X \rangle_{N_i}$ is the average of X using N_i samples. Both the fine approximation $Q^{f,i}$ and the coarse approximation $Q^{c,i}$ of the quantity Q at level i are random variables dependent on the same random input ω. The direct fine discretization and the two-scale model are used as follows:

$$\begin{aligned} Q^{f,i} &= Q_{h_i}\big(\mathcal{F}_{h_i,L}, \boldsymbol{K}_{h_i}\big), \\ Q^{c,i} &= Q_{H_i}\big(\mathcal{F}_{H_i,L}, P_{h_i}(\mathcal{F}_{h_i,H_i}, \boldsymbol{K}_{h_i})\big), \end{aligned} \qquad (12)$$

where $H_i = h_{i-1}$. All fracture sets $\mathcal{F}$ and fine tensor fields $\boldsymbol{K}_{h_i}$ are functions of a common input random state ω. To obtain a functional MLMC estimator, we must ensure that:

$$\mathbb{E}Q^{f,i} = \mathbb{E}Q^{c,i+1}, \qquad (13)$$

Discussion of this condition is beyond the scope of this paper; however, it is satisfied provided that the homogenized hydraulic conductivity $\boldsymbol{K}_{H_i}$ has the same distribution as $\boldsymbol{K}_{h_{i-1}}$ sampled directly. Regarding the requirements for the homogenization surrogates applicable to MLMC, we can make two key observations:

- The homogenization problems P_{h_i} only differ in the length scale determined by the block size $l_i = 1.5H_i$. The mesh resolution and size range of discrete fractures are the same, relative to l_i. Consequently, we may employ scale equivariance (Eq. ((10)) and evaluate the homogenized permeability field on the unit square: $\boldsymbol{K}_{H_i}(x) = l_i^2 P_h(x, 1, K_{h_i}/l_i^2)$.
- While the fracture hydraulic conductivity k_f scales with l_i^2 according to Eq. ((3)), this is not generally the case for the matrix hydraulic conductivity $\boldsymbol{K}_{h_i}$. Therefore, the ratio K_f/K_m varies between MLMC levels, necessitating a surrogate model that is robust to different hydraulic conductivity ratios.

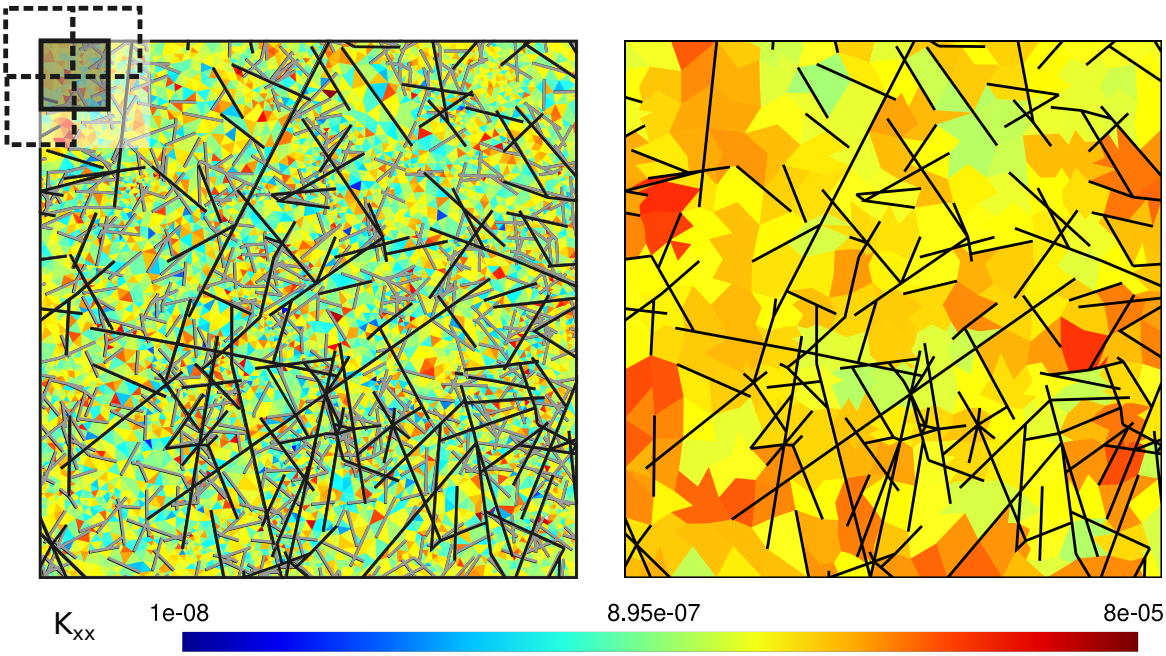


Fig. 1: **Left:** The fine-scale DFM model ($h < 5$), grey fractures, and permeability tensor field are homogenized using overlapping square blocks of size 15 depicted in the corner. **Right:** The coarse model ($H < 10$) with homogenized permeability tensor component K_{xx}. Notice the reduced spread of the upscaled field.

3 Surrogate architecture

Our deep learning (DL) surrogate solves a regression problem through supervised learning. It learns an unknown mapping from input random DFM parameters (SRF of matrix hydraulic conductivity tensors and fractures' hydraulic conductivity and cross section) to output equivalent tensors of hydraulic conductivity, utilizing collected sample pairs of inputs and corresponding outputs of a DFM model. The DL surrogate comprises a convolutional neural network (CNN) serving as a feature extractor and a feed-forward neural network (FNN) responsible for the regression task.

Convolutional neural networks [16, ch. 9] (CNNs) process data on regular grids, particularly images. Numerous architectures of CNNs (see [48]) have recently been developed for tasks such as image classification, segmentation, natural language processing, etc. CNNs utilize two fundamental operations: convolution and pooling. Popular software libraries Tensorflow and PyTorch use the term "convolution" to refer to cross-correlation. Consider a 2D image P of shape $I \times J \times C_{in}$, where C_{in} stands for the number of input channels, e.g., RGB images have three channels. Let R be a convolution kernel of shape $M \times N \times C_{in} \times C_{out}$, where C_{out} denotes the number of output channels. Then, the cross-correlation operation can be expressed as:

$$S(i,j,c_{out}) = (P * R)(i,j,c_{out}) = \sum_{m}\sum_{n}\sum_{c_{in}} P(i-m, j-n, c_{in}) R(m,n,c_{in},c_{out}). \tag{14}$$

Features extracted from the data (S) are passed through a non-linear activation function (e.g., ReLU), and pooling is often applied to enhance computational efficiency. A pooling layer usually reduces the dimensions of S by computing summary statistics (such as max, average, or sum) over small regions of each channel of S (e.g., 2×2, 4×4, etc.). Typically, the number of learnable parameters (i.e., kernel size) varies across convolutional layers. In particular, the number of channels often increases. Each channel can be thought of as responding to a different set of features, allowing

Table 1: The Architecture of the surrogate

Layer	Type	Output size
	Input	$256 \times 256 \times 4$
1	Conv2D	$127 \times 127 \times 24$
2	Conv2D	$62 \times 62 \times 48$
3	Conv2D	$30 \times 30 \times 96$
4	Conv2D	$14 \times 14 \times 192$
5	Conv2D	$6 \times 6 \times 256$
7	Linear	2048
8	Linear	2048
9	Linear	1024
10	LinearOutput	3

different channels to specialize in recognizing various objects, as described by Zhang et al. [49]. The key advantage of CNNs lies in parameter sharing. At each layer, the convolutional kernel R utilizes a single set of parameters shared across all locations. This characteristic contributes to the efficiency of CNNs compared to traditional fully connected neural networks (see [16, p. 328], for details).

We support both unstructured and structured computational meshes and utilize the Datashader software library (see [50]) for nearest-neighbor interpolation of potentially unstructured data into a 256×256 matrix. We consider tensors of hydraulic conductivity on each mesh element. For each pixel in the resulting matrix, we assign values of three independent components of the 2D tensor, along with one additional value representing the cross section of the element.

The CNN part of the surrogate consists of a Conv2D layer, which integrates convolution and max pooling. It employs a 3×3 convolutional kernel with a stride of 1 and zero padding. The kernel is accompanied by a 2×2 max pool with a stride of 2. We use batch normalization to facilitate the training process. After passing through 5 convolutional layers, the original input data, which has the shape of $256 \times 256 \times 4$, is reduced to the shape of $6 \times 6 \times 256$ and then flattened to enter the linear fully connected layers of 2048, 2048, and 1024 neurons, respectively. In these layers, as well as in the convolution layer, the ReLU nonlinear activation function is utilized. Finally, the output layer consists of 3 neurons with the identity activation function. The architecture of the surrogate is described in Table 1.

3.1 Metrics for surrogate accuracy assessment

Let $\mathcal{D} = \{(\mathbf{X}_j, \mathbf{K}_j)\}_{j=1}^{D}$ be a dataset of i.i.d. samples, where $\mathbf{X}_j \in \mathbb{R}^{4\times F}$ is a vector of input features (with F representing the number of pixels) and $\mathbf{K}_j \in \mathbb{R}^3$ is the vector of components k_{xx}, k_{xy}, and k_{yy} of the corresponding equivalent hydraulic conductivity tensor. Since supervised learning is carried out, $\mathcal{D}$ is divided into learning (=training) samples $\mathcal{L}$, validation samples $\mathcal{V}$, and test samples $\mathcal{T}$. Given $\mathcal{L}$, the surrogate learns a function $f(\mathbf{x}) : \mathbb{R}^{|\mathbf{X}|} \to \mathbb{R}^3$. During the training process, the accuracy of the surrogate's predictions is evaluated by a loss function, in our case, the mean squared error (MSE): $\gamma(\mathcal{M}, f) = \frac{1}{|\mathcal{M}|}\|\mathbf{K}_\mathcal{M} - f(\mathbf{X}_\mathcal{M})\|_2^2$, where $\mathcal{M} \subseteq \mathcal{D}$. The training loss $\gamma(\mathcal{L}, f)$ and validation loss $\gamma(\mathcal{V}, f)$ are observed.

For the sake of comparing the accuracy of surrogates, we utilize:

- the coefficient of determination for component i of $\mathbf{K}$:

$$R_i^2 = 1 - \frac{\sum_\mathcal{M} (\mathbf{K}^i - f(\mathbf{X}))^2}{\sum_\mathcal{M} (\mathbf{K}^i - \overline{\mathbf{K}^i_\mathcal{M}})^2}, \qquad (15)$$

  where $\overline{\mathbf{K}^i_\mathcal{M}}$ is the average of $\mathbf{K}^i$ over dataset $\mathcal{M}$. For further analysis, we also employ $\overline{R^2} = \frac{R^2_{xx}+R^2_{xy}+R^2_{yy}}{3}$. The closer R^2 is to 1, the better.
- normalized root mean squared error for component i of $\mathbf{K}$:

$$NRMSE_i = \frac{\sqrt{\frac{1}{|\mathcal{M}|}\|\mathbf{K}^i_\mathcal{M} - f(\mathbf{X}_\mathcal{M})\|_2^2}}{\mathrm{std}(\mathbf{K}^i_\mathcal{M})}, \qquad (16)$$

  where $\mathrm{std}(\mathbf{K}^i_\mathcal{M})$ is the standard deviation of $\mathbf{K}^i$ over the dataset $\mathcal{M}$. The lower the *NRMSE*, the better. Moreover, if the *NRMSE* is above 1, we can use a simple random generator instead of a complex surrogate.

4 Dataset configuration

Datasets are constructed following our homogenization procedure described in Section 2.4. The original domain $\Omega_o = (0, 100)^2$ is extended to the domain $\Omega = (0, 114.28)^2$. A DFN is generated with the density $\rho'_{2D} = 10.0$, comprising approximately 1500 fractures. Fracture centers are placed

in Ω by the Poisson process. We consider a power-law distribution of fracture lengths ranging from $\underline{r} = 4.325$ to $\overline{r} = 100.0$, and the orientation angle is uniformly distributed. Based on this, the geometry with a mesh with step $h = 4.325$ is generated. The extended domain Ω is covered by 225 overlapping homogenization blocks of size $(14.28, 14.28)$, where the centers of corner homogenization blocks are placed in the corners of the original domain Ω_o. The matrix SRF is generated with parameters $\boldsymbol{\mu} = [-6. - 5.8]$, $\boldsymbol{\Sigma} = \begin{bmatrix} 0.25 & 0.2 \\ 0.2 & 0.25 \end{bmatrix}$.

We have analyzed the problem for diverse fracture-to-matrix hydraulic conductivity ratios K_f/K_m, where $K_f/K_m \in \{1 \times 10^3, 1 \times 10^5, 1 \times 10^7\}$ was selected as representative enough for this study. Consequently, three datasets with different K_f/K_m ratios of inputs were formed:

- **Dataset** $\mathcal{A}$: $K_f/K_m = 1 \times 10^3$,
- **Dataset** $\mathcal{B}$: $K_f/K_m = 1 \times 10^5$,
- **Dataset** $\mathcal{C}$: $K_f/K_m = 1 \times 10^7$.

Each dataset contains $75,000$ samples. In order to provide predictions for a range of matrix SRF correlation lengths up to $\lambda = 25$, correlation lengths $\lambda \in \{0, 10, 25\}$ are equally represented in the datasets. The $75,000$ samples are split into 80% training samples and 20% test samples. Within the training samples, 80% are allocated for actual training, while 20% are reserved for validation.

Distributions of equivalent hydraulic conductivity tensors for datasets $\mathcal{A}$,$\mathcal{B}$, and $\mathcal{C}$ are depicted in Fig.2. Considering k_{xx} and k_{yy}, we observe that for Dataset $\mathcal{A}$, the equivalent tensor is predominantly determined by matrix elements. For Dataset $\mathcal{B}$, we note an increased impact of fractures. However, there is still no visible percolation threshold characterized by a sharp increase in equivalent hydraulic conductivity at the moment of significant interconnection of fractures. This phenomenon has been observed in studies such as [13, 51]. As the ratio K_f/K_m increases, the gap between non-percolated and percolated samples widens. We observe orders of magnitude difference for Dataset $\mathcal{C}$. Regarding k_{xy}, the shape of the data is very similar across datasets.

The equivalent tensor distributions exhibit significant dissimilarities in scale and shape. Consequently, training a surrogate model on the combined dataset of all presented K_f/K_m ratios has proven to be extremely challenging and has resulted in unsatisfactory accuracy thus far. Therefore, we resort to training separate surrogate models: Surrogate A for Dataset $\mathcal{A}$, Surrogate B for Dataset $\mathcal{B}$, and Surrogate C for Dataset $\mathcal{C}$. Subsequently, based on the K_f/K_m ratio of a particular input $\mathbf{X}_j$, the most suitable surrogate is selected for prediction.

4.1 Dataset preprocessing

In order to facilitate the surrogate's training, hydraulic conductivity tensors in the dataset inputs $\mathbf{X}$ and outputs $\mathbf{K}$ are preprocessed in the same way. For each sample j, the average matrix hydraulic conductivity $\overline{\mathbf{X}^j}_m$ (average over the entire tensor) is calculated.

First, inputs and outputs are normalized so that hydraulic conductivities of different scales with the same K_f/K_m value can be accommodated:

$$\begin{aligned} \mathbf{X}^j &= \mathbf{X}^j / \overline{\mathbf{X}^j}_m \\ \mathbf{K}_j &= \mathbf{K}_j / \overline{\mathbf{X}^j}_m \end{aligned} \tag{17}$$

Second, data is independently standardized for input and output hydraulic conductivities, where mean (avg) and standard deviation (std) are calculated based on the training set $\mathcal{L}$. Each component of a hydraulic conductivity tensor is preprocessed as follows:

$$\tilde{k}^j_{xx} = \frac{\log(k^j_{xx}) - \mathrm{avg}_{\mathcal{L}}(\log(k_{xx}))}{\mathrm{std}_{\mathcal{L}}(\log(k_{xx}))},$$

$$\tilde{k}^j_{xy} = \frac{k^j_{xy} - \mathrm{avg}_{\mathcal{L}}(k_{xy})}{\mathrm{std}_{\mathcal{L}}(k_{xy})},$$

$$\tilde{k}^j_{yy} = \frac{\log(k^j_{yy}) - \mathrm{avg}_{\mathcal{L}}(\log(k_{yy}))}{\mathrm{std}_{\mathcal{L}}(\log(k_{yy}))}.$$

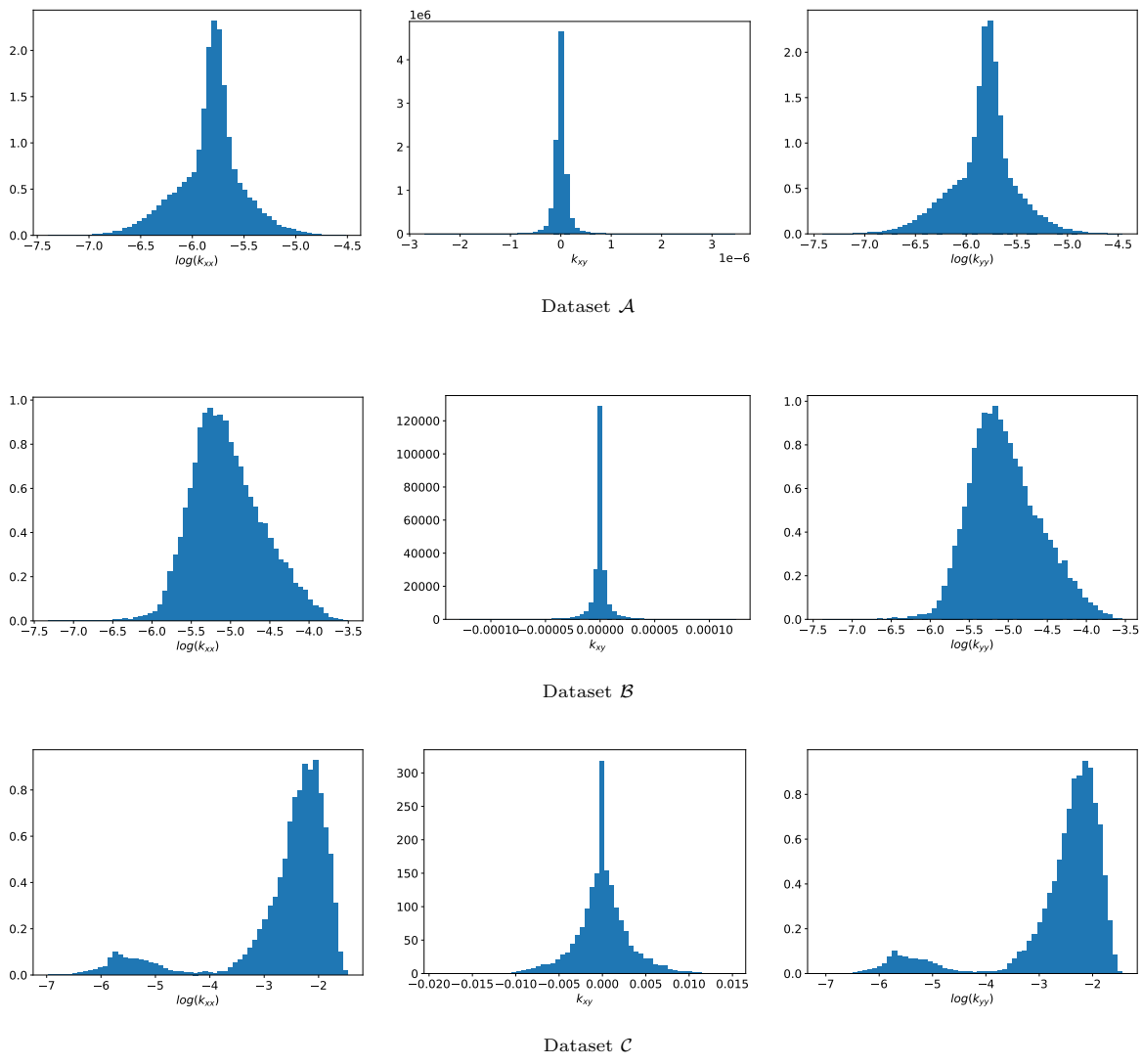


Fig. 2: Distributions of $\boldsymbol{K}^{eq}$ components for datasets of different K_f/K_m, Dataset $\mathcal{A}$ for $K_f/K_m = 1 \times 10^3$, Dataset $\mathcal{B}$ for $K_f/K_m = 1 \times 10^5$, and Dataset $\mathcal{C}$ for $K_f/K_m = 1 \times 10^7$.

5 Results

In this section, we present the prediction accuracy of trained surrogates and discuss their prediction capabilities on datasets with different matrix SRF correlation lengths λ and fracture densities ρ'_{2D}. Additionally, we employ trained surrogates for upscaling a DFM model within two macroscale problems to demonstrate the impact of prediction accuracy of $\boldsymbol{K}^{eq}$ on a coarser scale. We use 125 epochs for surrogate training, and the model with the lowest validation loss is saved for further use. During training, the initial learning rate $\alpha = 0.0025$ adaptively decreases to 10 percent of the previous value when the validation loss does not decrease in 10 subsequent epochs.

Fig.3 shows the targets-predictions plot along with R^2 and NRMSE values for the test sets. The best results ($R^2 > 0.95$) were achieved for Surrogate A, where the distribution of tensor components is relatively narrow. These results are comparable to or better than those provided by Pal et al. [27], who do not consider fractures, or

by Zhu et al. [52], who only consider DFN models, meaning the impact of the matrix on the equivalent hydraulic conductivity is not considered.

As the training data distribution becomes broader and more complex (Fig. 2), the surrogate accuracy deteriorates (Fig. 3). This is particularly visible for Surrogate C, where the bimodal distributions of $\log(k_{xx})$ and $\log(k_{yy})$ lead to poor accuracy for samples between two distant peaks. This is due to the lack of this specific data in the training set compared to the data in the distribution peaks (Fig. 2), especially the larger peak, where the predictions tend to be more accurate. This effect intensifies with increasing K_f/K_m. For all surrogates, the off-diagonal component of $\boldsymbol{K}^{eq}$ is predicted with higher error than the diagonal components. Due to the nature of k_{xy}, we cannot carry out logarithmic-like preprocessing. Consequently, k_{xy} suffers from a scarcity of training data in tail areas, where a more pronounced dispersion of predictions is observed.

Although using a larger training dataset generally improves the prediction accuracy of deep neural networks, it may not effectively address the issues associated with underrepresented data if the distribution of the additional training data remains the same. Instead of simply increasing the size of the training set, we argue that to enhance prediction accuracy, one should focus on generating more data in the tails for Surrogate B or in the area between the lower and higher peaks for Surrogate C. This process is challenging because we lack a straightforward method to generate DFM homogenization blocks for a given equivalent hydraulic conductivity tensor.

In the following subsections, we investigate the prediction accuracy of trained surrogates on datasets with different fracture densities ρ'_{2D} and on datasets of different correlation lengths λ of matrix SRF. Subsequently, we discuss the speedup achieved by surrogates.

5.0.1 Impact of fracture density

The impact of fracture density on surrogate accuracy has been previously discussed by Zhu et al. [52], where they predicted equivalent permeability tensors based on DFN models. They generally observed a deterioration in prediction accuracy with increasing fracture density, which aligns with our findings. Fig. 4 illustrates prediction accuracy for $\rho'_{2D} \in \{2.5, 5, 7.5, 10, 12.5, 15\}$. Surrogates trained with $\rho'_{2D} = 10$ achieve increasingly better accuracy when predicting on datasets with $\rho'_{2D} < 10$. Since $\rho'_{2D} = 10$ is prescribed for the entire domain Ω, individual homogenization blocks may contain a number of fractures that corresponds to $\rho'_{2D} < 10$. However, for $\rho'_{2D} > 10$, $\overline{R^2}$ diminishes, with the most significant reduction observed in the case of Surrogate C, where fractures play a dominant role in $\boldsymbol{K}^{eq}$ calculation. Still, $\overline{R^2} \approx 0.68$ might suffice for some applications. If not, surrogate training on higher ρ'_{2D} might be a viable solution. It's important to note that even for Surrogate C trained on $\rho'_{2D} = 10$ data, some k_{xy} components are inaccurately predicted for $\rho'_{2D} \in \{2.5, 5\}$. The reason is that the k_{xy} values span 7 orders of magnitude (1×10^{-9} to 1×10^{-2}). Given the preprocessing of k_{xy} and loss function, larger k_{xy} values are predicted more accurately. When ρ'_{2D} is small enough, there are instances where the off-diagonal component of the hydraulic conductivity tensor is among the smallest in the dataset (around 1×10^{-8} and lower). This leads to substantial relative errors in predictions and, consequently, potentially non-positive definite predicted $\boldsymbol{K}^{eq}$ tensors. We do not encounter this issue for $K_f/K_m \leq 1 \times 10^5$, which is the maximum ratio investigated in studies on equivalent permeability/hydraulic conductivity calculation, such as [10, 13, 31, 52].

5.0.2 Impact of correlation length

We further investigate the impact of the matrix's SRF correlation length λ on the prediction accuracy of our surrogates trained on $\lambda \in \{0, 10, 25\}$. As shown in Fig. 5, Surrogate A exhibits improved prediction accuracy with increasing λ. This trend is expected because a larger correlation length induces lower variability in the SRF field, making it easier for neural networks to learn. The opposite trend is observed in the cases of Surrogate B and Surrogate C. As mentioned earlier, these surrogates were trained on datasets with $K_f/K_m = 1 \times 10^5$ and $K_f/K_m = 1 \times 10^7$, where the hydraulic conductivity of fractures significantly influences $\boldsymbol{K}^{eq}$. Consequently, samples with very similar DFNs but dissimilar SRFs can result in similar $\boldsymbol{K}^{eq}$. As the correlation length increases, it becomes increasingly

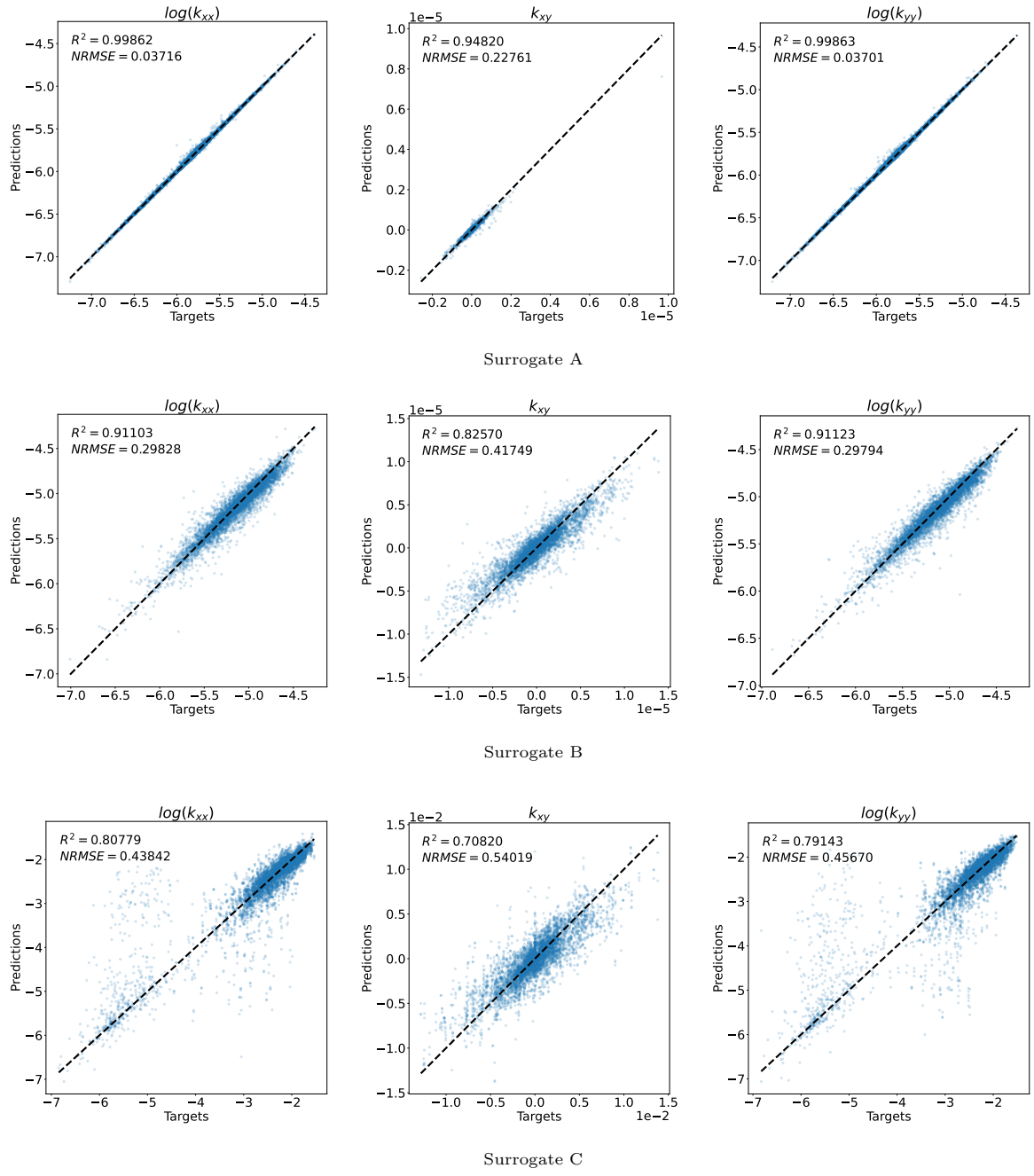


Fig. 3: Accuracy of the surrogates on test datasets

challenging for Surrogate B and Surrogate C to differentiate between these two cases, potentially leading to lower prediction accuracy. Despite the equal representation of all three correlation lengths $\lambda = \{0, 10, 25\}$ in the training dataset, Surrogate B and Surrogate C exhibit the best predictive performance for uncorrelated SRFs ($\lambda = 0$).

Another important observation is that although correlation lengths $\lambda = \{5, 15\}$ were

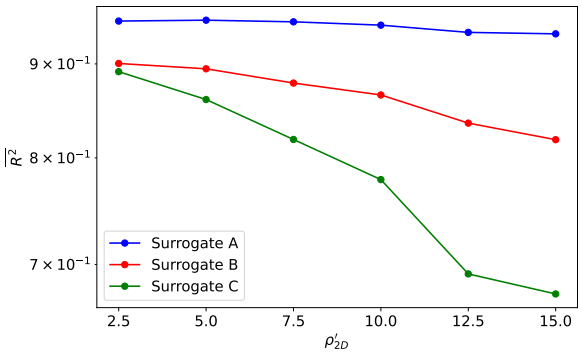


Fig. 4: Impact of fracture density ρ'_{2D} on surrogate's accuracy

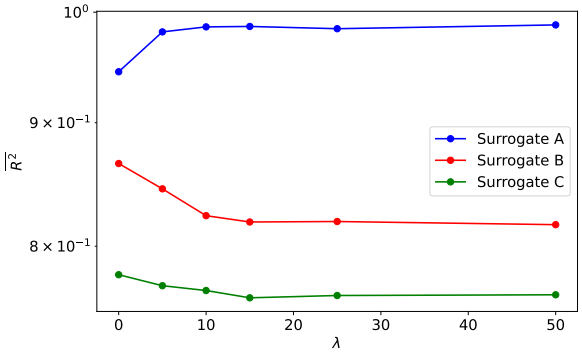


Fig. 5: Impact of matrix correlation length λ on surrogate's accuracy

not explicitly included in the training datasets, the corresponding predictive performance mirrors the trends observed for the included correlation lengths. This suggests that a wide range of correlation lengths on the homogenization block (of size 14.24×14.28) can be effectively covered by using datasets with $\lambda = \{0, 10, 25\}$, where the SRFs are generated following the procedure described in Section 2.2.

5.0.3 Computational cost reduction

We aim to utilize surrogates to speed up numerical homogenization. To this end, we compare the computational cost (CPU time) required for numerical homogenization (C_H) and surrogate prediction (C_S) across varying numbers of homogenization blocks covering the entire domain

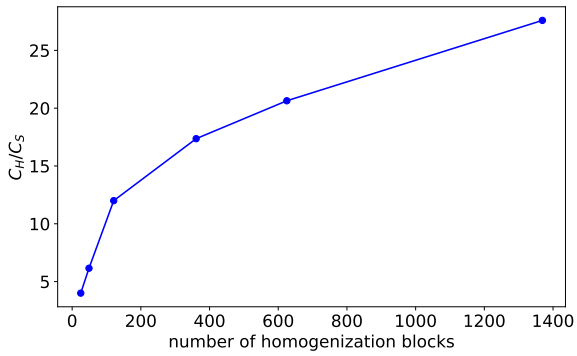


Fig. 6: Speedup gained by surrogate predictions for different numbers of homogenization blocks

Ω (see Fig. 6). In the case of C_S, the predominant cost arises from the rasterization of an input mesh. The entire domain Ω is rasterized once and then divided into homogenization blocks. The prediction time of the neural network is negligible, and training time is excluded from consideration. The cost of numerical homogenization C_H encompasses the generation of a computational mesh and the execution of simulation software for each homogenization block. Consequently, the acceleration achieved by using the surrogates increases with the number of homogenization blocks. For instance, with 25 blocks, $C_H/C_S = 4$, while with 1369 blocks, we attain $C_H/C_S \approx 28$. The latter scenario would correspond to the finest-scale samples, where we also assume the most accurate surrogate predictions.

5.1 Upscaled equivalent hydraulic conductivity tensors in use

We explore the impact of utilizing predicted equivalent tensors for the Aquifer problem and the Anisotropy problem, as described in Section 2.3. In both cases, the effect of upscaling by numerical homogenization and by surrogates is compared. We consider fine DFM samples with $\rho'_{2D} = 10$, $\lambda = 0$, $K_f/K_m \in \{1 \times 10^3, 1 \times 10^5, 1 \times 10^7\}$. Trained Surrogate A, Surrogate B, and Surrogate C are used for predicting $\boldsymbol{K}^{eq}$. In accordance with the procedure explained in Section 2.4, homogenized/predicted $\boldsymbol{K}^{eq}$ are utilized for the matrix hydraulic conductivity of the upscaled DFM model.

Table 2: R^2 comparison of components of calculated equivalent hydraulic conductivity tensors $\boldsymbol{K}^{eq}$ of DFM models upscaled by numerical homogenization and by surrogate

R^2	k_{xx}	k_{xy}	k_{yy}
Surrogate A	0.99967	0.99990	0.99910
Surrogate B	0.99960	0.99984	0.99989
Surrogate C	0.99998	0.99998	0.99998

5.1.1 Aquifer problem

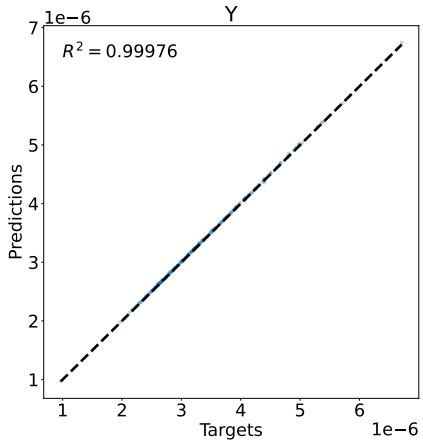
First, we compare upscaled DFM models on the Aquifer problem described in Section 2.3.1. Fig. 7 presents the targets-predictions plot and R^2 for variable Y. For $K_f/K_m = 1 \times 10^3$, Surrogate A's predictions lead to almost the same Y as using the numerical homogenization. In the case of $K_f/K_m = 1 \times 10^5$, Surrogate B's lower prediction accuracy causes a slightly lower $R^2 \approx 0.95$ for Y. Naturally, for $K_f/K_m = 1 \times 10^7$, the deteriorated prediction accuracy of Surrogate C leads to a somewhat smaller R^2 of about 0.87 for Y as well.

While in the case of $K_f/K_m = 1 \times 10^3$, the whole upscaled SRF plays an equally important role in calculating Y (see pressure fields in Fig. 8), in the case of $K_f/K_m = 1 \times 10^7$ the influence of fractures is so dominant that mainly upscaled hydraulic conductivities near the left and right vertical boundaries are significant (see pressure fields in Fig. 9). Therefore, the latter case is more sensitive to $\boldsymbol{K}^{eq}$ prediction accuracy on considerably fewer mesh elements.

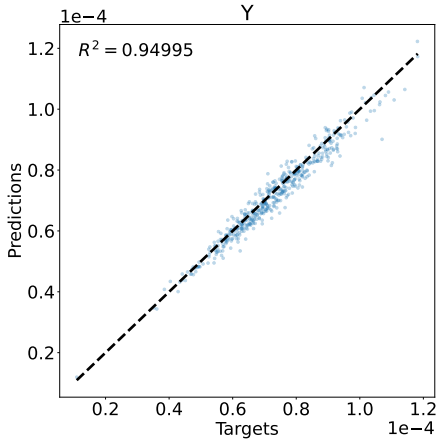
5.1.2 Anisotropy problem

Second, we compare equivalent hydraulic conductivity tensors for DFM models upscaled by numerical homogenization and by our surrogates. For comparison, we calculate $\boldsymbol{K}^{eq}$ as described in Section 2.3.2 but for the entire domain Ω_0 of the upscaled DFM model. The obtained R^2 (see Table 2) shows that all presented surrogates provide very accurate predictions of $\boldsymbol{K}^{eq}$, regardless of the value of $K_f/K_m \in \{1 \times 10^3, 1 \times 10^5, 1 \times 10^7\}$ used. In other words, the components of $\boldsymbol{K}^{eq}$ of upscaled DFM models are nearly identical for numerical homogenization and surrogate modeling.

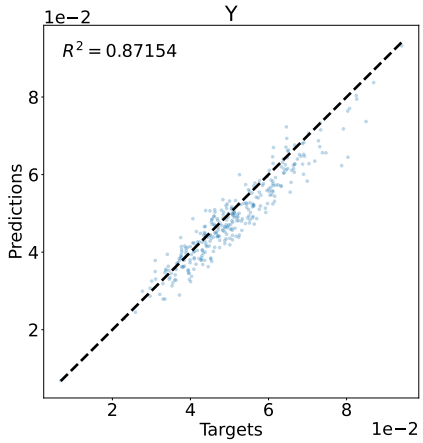
In Fig. 10, SRFs of k_{xx} for an upscaled DFM model with $K_f/K_m = 1 \times 10^5$ are depicted. For the sake of clarity, hydraulic conductivity (K_xx) scales are adjusted to have the same range for



(a) Surrogate A

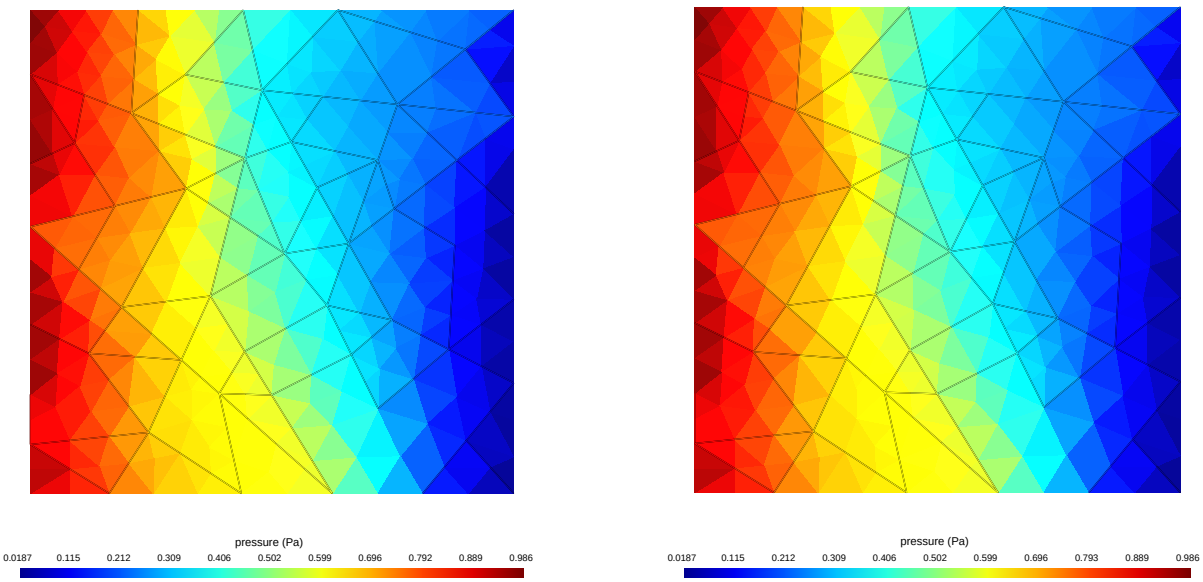


(b) Surrogate B



(c) Surrogate C

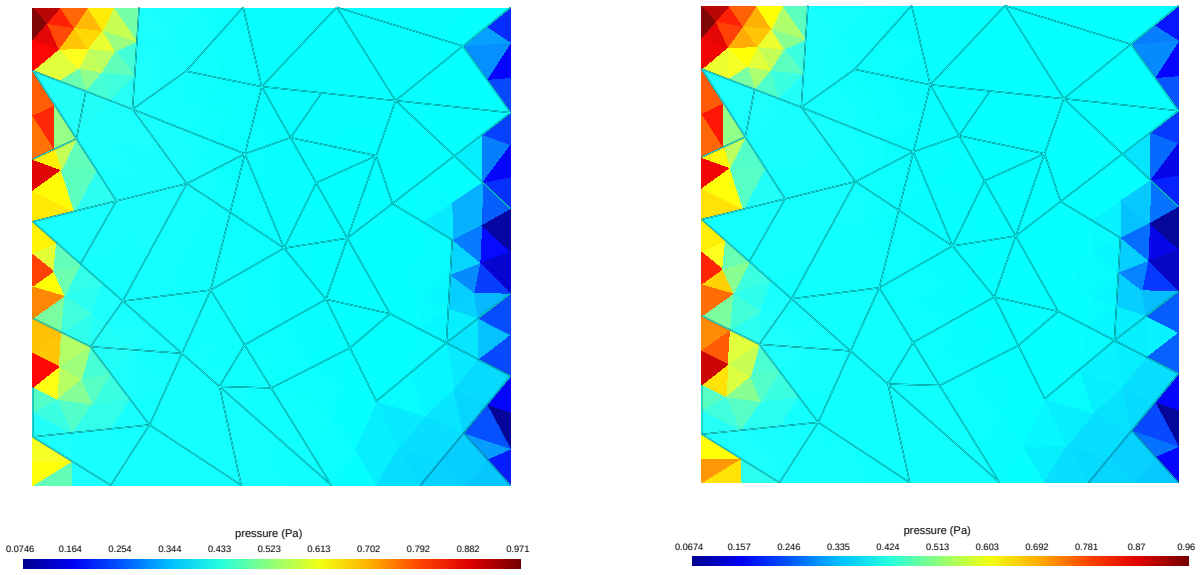
Fig. 7: Comparison of outflow Y from DFM models upscaled by numerical homogenization and by surrogates, 1000 samples.



(a) Upscaled DFM model by numerical homogenization

(b) Upscaled DFM model by Surrogate A

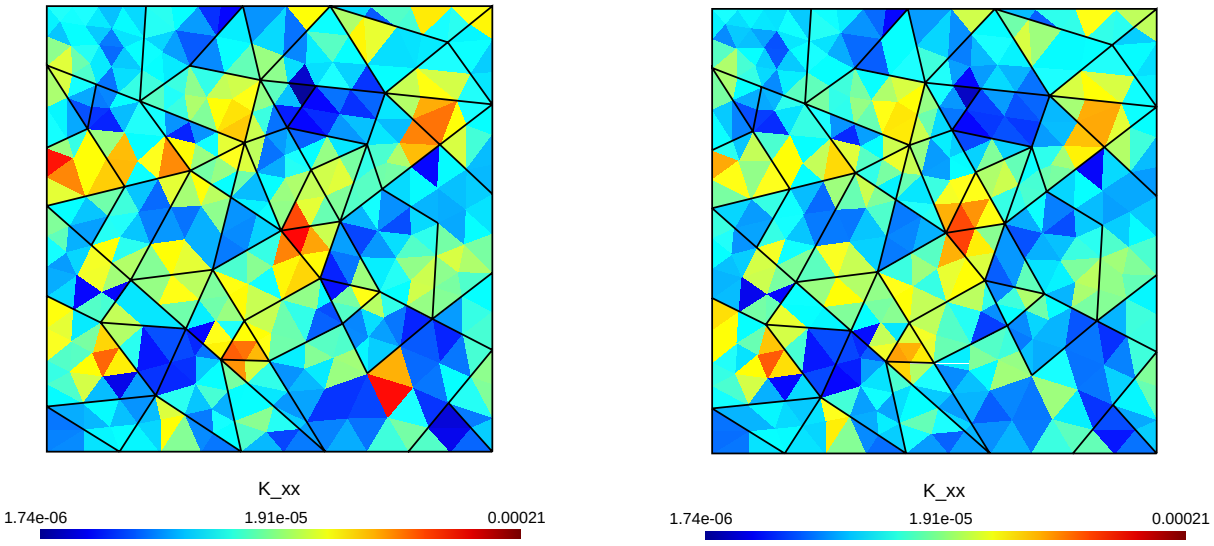
Fig. 8: Aquifer problem - pressure field example, $K_f/K_m = 1 \times 10^3$



(a) Upscaled DFM model by numerical homogenization

(b) Upscaled DFM model by Surrogate C

Fig. 9: Aquifer problem - pressure field example, $K_f/K_m = 1 \times 10^7$



(a) Upscaled SRF of k_{xx} by numerical homogenization

(b) Upscaled SRF of k_{xx} by Surrogate B

Fig. 10: Upscaled SRF of k_{xx} by numerical homogenization and by Surrogate B, $K_f/K_m = 1 \times 10^5$

upscaling by Surrogate B and upscaling by numerical homogenization. The hydraulic conductivities of fractures are left out. Despite some dissimilarities, especially for higher values of k_{xx}, $\boldsymbol{K}^{eq}$ tensors are extremely similar. The explanation is that as the predictive accuracy of our surrogates decreases with increasing K_f/K_m, the impact of fracture conductivities increases simultaneously. Overall, we have highly accurate Surrogate A for $K_f/K_m = 1 \times 10^3$. Thus, $\boldsymbol{K}^{eq}$ is also highly accurate. For $K_f/K_m \in \{1 \times 10^5, 1 \times 10^7\}$, the inaccuracy in predictions is compensated by the increased influence of fractures.

6 Conclusions

This article presented the deep learning-based surrogate for predicting the equivalent hydraulic conductivity tensor $\boldsymbol{K}^{eq}$ resulting from the upscaling of 2D DFM models. In addition to assessing the prediction accuracy of surrogates, we discussed their limitations and demonstrated practical usage on two macroscale problems.

We investigated different fracture-to-matrix hydraulic conductivity K_f/K_m ratios and formed three separate datasets of the same structure for $K_f/K_m \in \{1\times10^3, 1\times10^5, 1\times10^7\}$. For each ratio, a separate surrogate of the same architecture was trained and evaluated on test datasets. Surrogates achieved highly accurate ($R^2 > 0.95$) predictions for all three independent components of $\boldsymbol{K}^{eq}$ when $K_f/K_m \leq 10^3$. However, prediction accuracy deteriorated with increasing K_f/K_m due to the complex distribution of $\boldsymbol{K}^{eq}$ components. The off-diagonal component of $\boldsymbol{K}^{eq}$ exhibited consistently lower accuracy than diagonal components across all the cases. This inaccuracy is mainly driven by higher errors of predictions for data in the tails of the distributions. The speedup gained by surrogates depended on the number of homogenization blocks. From 4× for 25 blocks to 28× for 1396 blocks used for the upscaling of the smallest-scale DFM model in the study.

We also compared prediction accuracy for different fracture densities ρ'_{2D} and correlation lengths λ of matrix SRF. In accordance with the literature, we observed improved accuracy on lower ρ'_{2D} than was explicitly used for training. Regarding correlation lengths, we noticed an improvement in prediction accuracy with increasing λ for the case of lower impact of fractures ($K_f/K_m = 1\times10^3$) and conversely, a deterioration in prediction accuracy with increasing λ for cases with the higher impact of fractures ($K_f/K_m \in \{1\times10^5, 1\times10^7\}$). In all the cases, our surrogates exhibited expectedly accurate predictions even for λ that was not included in the training set, up to

$\lambda = 50$. In this matter, our surrogates generalize well.

Although the higher the prediction accuracy, the better, the necessary level of accuracy is not always obvious. Therefore, we investigated the impact of predicted $\boldsymbol{K}^{eq}$ on two macroscale problems. In the case of comparing the usage of surrogates to numerical homogenization for upscaling of DFM models for the Aquifer problem, we observed $R^2 \gtrapprox 0.95$ for $K_f/K_m \in \{1 \times 10^3, 1 \times 10^5\}$. Even the observed $R^2 \approx 0.87$ for $K_f/K_m = 1 \times 10^7$ might be sufficient. Such extreme hydraulic conductivity ratios may appear on coarse levels of MLMC, where the requested accuracy of the model is smaller. Regarding the Anisotropy problem, the prediction capability of all of the presented surrogates was excellent ($R^2 \gtrapprox 0.999$) for upscaled DFM models.

We are also aware of the limitations of our approach and recognize the opportunities for improvement. In rare cases where $K_f/K_m > 1 \times 10^7$ and $\rho'_{2D} \leq 5$, prediction accuracy could be enhanced. Possible strategies include augmenting the training dataset in regions of smaller peaks and between peaks, as well as preprocessing k_{xy} differently to ensure accurate prediction across all scales of values. For other considered K_f/K_m ratios ($< 1 \times 10^7$), our surrogates demonstrated sufficient performance sufficiently for use within a multiscale MLMC scheme. This will be investigated in future work along with extending our approach to 3D DFM models.

Acknowledgements. The research was supported within the framework of EURAD, the European Joint Programme on Radioactive Waste Management (Grant Agreement No 847593). This work was (partly) supported by the Student Grant Scheme at the Technical University of Liberec through project nr. SGS-2023-3379. Computational resources were provided by the e-INFRA CZ project (ID:90254), supported by the Ministry of Education, Youth and Sports of the Czech Republic.

Code availability The source codes are available for downloading at the link: https://github.com/martinspetlik/MLMC-DFM/

Author contribution Martin Špetlík: Writing - original draft, Software, Experimentation. Jan Březina: Supervision, Writing - review & editing, Software. Eric Laloy: Supervision, Writing - review & editing.

References

[1] Banks, D., Robins, N.: An Introduction to Groundwater in Crystalline Bedrock. Geological Survey of Norway, Trondheim, Norway (2002)

[2] Giles, M.B.: Multilevel Monte Carlo methods. Acta Numerica **24**, 259–328 (2015) https://doi.org/10.1017/S096249291500001X

[3] Auriault, J.-L., Boutin, C., Geindreau, C.: Homogenization of Coupled Phenomena in Heterogenous Media. ISTE, London, UK (2009). https://doi.org/10.1002/9780470612033

[4] Bonnet, E., Bour, O., Odling, N.E., Davy, P., Main, I., Cowie, P., Berkowitz, B.: Scaling of fracture systems in geological media. Reviews of Geophysics **39**(3), 347–383 (2001) https://doi.org/10.1029/1999RG000074

[5] Sanchez-Vila, X., Guadagnini, A., Carrera, J.: Representative hydraulic conductivities in saturated groundwater flow. Reviews of Geophysics **44**(3) (2006) https://doi.org/10.1029/2005RG000169

[6] Renard, P., de Marsily, G.: Calculating equivalent permeability: a review. Advances in Water Resources **20**(5), 253–278 (1997) https://doi.org/10.1016/S0309-1708(96)00050-4

[7] Chen, T., Clauser, C., Marquart, G., Willbrand, K., Mottaghy, D.: A new upscaling method for fractured porous media. Advances in Water Resources **80**, 60–68 (2015) https://doi.org/10.1016/j.advwatres.2015.03.009

[8] Farmer, C.L.: Upscaling: a review. International Journal for Numerical Methods in Fluids **40**(1-2), 63–78 (2002) https://doi.org/10.1002/fld.267

[9] Renard, P., Ababou, R.: Equivalent permeability tensor of heterogeneous media:

Upscaling methods and criteria (review and analyses). Geosciences **12**(7) (2022) https://doi.org/10.3390/geosciences12070269

[10] Bogdanov, I.I., Mourzenko, V.V., Thovert, J.-F., Adler, P.M.: Effective permeability of fractured porous media in steady state flow. Water Resources Research **39**(1) (2003) https://doi.org/10.1029/2001WR000756

[11] Koudina, N., Garcia, R.G., Thovert, J.-F., Adler, P.M.: Permeability of three-dimensional fracture networks. Physical Review E **57**(4), 4466–4479 (1998) https://doi.org/10.1103/PhysRevE.57.4466

[12] Bogdanov, I.I., Mourzenko, V.V., Thovert, J.-F., Adler, P.M.: Effective permeability of fractured porous media with power-law distribution of fracture sizes. Phys. Rev. E **76**, 036309 (2007) https://doi.org/10.1103/PhysRevE.76.036309

[13] Lang, P.S., Paluszny, A., Zimmerman, R.W.: Permeability tensor of three-dimensional fractured porous rock and a comparison to trace map predictions. Journal of Geophysical Research: Solid Earth **119**(8), 6288–6307 (2014) https://doi.org/10.1002/2014JB011027

[14] Lee, S.H., Lough, M.F., Jensen, C.L.: Hierarchical modeling of flow in naturally fractured formations with multiple length scales. Water Resources Research **37**(3), 443–455 (2001) https://doi.org/10.1029/2000WR900340

[15] Azizmohammadi, S., Matthäi, S.K.: Is the permeability of naturally fractured rocks scale dependent? Water Resources Research **53**(9), 8041–8063 (2017) https://doi.org/10.1002/2016WR019764

[16] Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge, Massachusetts (2016). http://www.deeplearningbook.org

[17] Caglar, B., Broggi, G., Ali, M.A., Orgéas, L., Michaud, V.: Deep learning accelerated prediction of the permeability of fibrous microstructures. Composites Part A: Applied Science and Manufacturing **158**, 106973 (2022) https://doi.org/10.1016/j.compositesa.2022.106973

[18] Alqahtani, N., Alzubaidi, F., Armstrong, R.T., Swietojanski, P., Mostaghimi, P.: Machine learning for predicting properties of porous media from 2d x-ray images. Journal of Petroleum Science and Engineering **184**, 106514 (2020) https://doi.org/10.1016/j.petrol.2019.106514

[19] Wang, Y., Li, H., Xu, J., Liu, S., Wang, X.: Machine learning assisted relative permeability upscaling for uncertainty quantification. Energy **245**, 123284 (2022) https://doi.org/10.1016/j.energy.2022.123284

[20] Rao, C., Liu, Y.: Three-dimensional convolutional neural network (3d-cnn) for heterogeneous material homogenization. Computational Materials Science **184**, 109850 (2020) https://doi.org/10.1016/j.commatsci.2020.109850

[21] Hong, J., Liu, J.: Rapid estimation of permeability from digital rock using 3d convolutional neural network. Computational Geosciences **24**(4), 1523–1539 (2020) https://doi.org/10.1007/s10596-020-09941-w

[22] Peng, H., Liu, A., Huang, J., Cao, L., Liu, J., Lu, L.: Ph-net: Parallelepiped microstructure homogenization via 3d convolutional neural networks. Additive Manufacturing **60**, 103237 (2022) https://doi.org/10.1016/j.addma.2022.103237

[23] Vasilyeva, M., Tyrylgin, A.: Machine learning for accelerating macroscopic parameters prediction for poroelasticity problem in stochastic media. Computers & Mathematics with Applications **84**, 185–202 (2021) https://doi.org/10.1016/j.camwa.2020.09.024

[24] Liu, M., Ahmad, R., Cai, W., Mukerji, T.: Hierarchical homogenization with deep-learning-based surrogate model for rapid estimation of effective permeability from digital rocks. Journal of Geophysical Research: Solid Earth **128**(2), 2022–025378 (2023) https://doi.org/10.1029/2022JB025378

[25] Meng, Y., Jiang, J., Wu, J., Wang, D.: Transformer-based deep learning models for predicting permeability of porous media. Advances in Water Resources **179**, 104520 (2023) https://doi.org/10.1016/j.advwatres.2023.104520

[26] Stepanov, S., Spiridonov, D., Mai, T.: Prediction of numerical homogenization using deep learning for the richards equation. Journal of Computational and Applied Mathematics **424**, 114980 (2023) https://doi.org/10.1016/j.cam.2022.114980

[27] Pal, M., Makauskas, P., Malik, S.: Upscaling porous media using neural networks: A deep learning approach to homogenization and averaging. Processes **11**(2) (2023) https://doi.org/10.3390/pr11020601

[28] Cai, C., Vlassis, N., Magee, L., Ma, R., Xiong, Z., Bahmani, B., Wong, T.-F., Wang, Y., Sun, W.: Equivariant geometric learning for digital rock physics: estimating formation factor and effective permeability tensors from morse graph. International Journal for Multiscale Computational Engineering **21**(5), 1–24 (2023)

[29] Ferreira, C.A.S., Kadeethum, T., Bouklas, N., Nick, H.M.: A framework for upscaling and modelling fluid flow for discrete fractures using conditional generative adversarial networks. Advances in Water Resources **166**, 104264 (2022) https://doi.org/10.1016/j.advwatres.2022.104264

[30] He, X., Santoso, R., Alsinan, M., Kwak, H., Hoteit, H.: Constructing Dual-Porosity Models from High-Resolution Discrete-Fracture Models Using Deep Neural Networks. SPE Reservoir Simulation Conference, vol. Day 1 Tue, October 26, 2021, pp. 011–014012 (2021). https://doi.org/10.2118/203901-MS

[31] Andrianov, N.: Upscaling of two-phase discrete fracture simulations using a convolutional neural network. Computational Geosciences **26**, 1237–1259 (2022) https://doi.org/10.1007/s10596-022-10149-3

[32] Berre, I., Doster, F., Keilegavlen, E.: Flow in fractured porous media. Transport in Porous Media **130**(1), 215–236 (2019) https://doi.org/10.1007/s11242-018-1171-6

[33] Long, J.C.S., Remer, J.S., Wilson, C.R., Witherspoon, P.A.: Porous media equivalents for networks of discontinuous fractures. Water Resources Research **18**(3), 645–658 (1982) https://doi.org/10.1029/WR018i003p00645

[34] Bour, O., Davy, P.: Connectivity of random fault networks following a power law fault length distribution. Water Resources Research **33**(7), 1567–1583 (1997) https://doi.org/10.1029/96WR00433

[35] de Dreuzy, J.-R., Méheust, Y., Pichot, G.: Influence of fracture scale heterogeneity on the flow properties of three-dimensional discrete fracture networks (DFN). Journal of Geophysical Research: Solid Earth **117**(B11) (2012) https://doi.org/10.1029/2012JB009461

[36] Hardebol, N.J., Maier, C., Nick, H., Geiger, S., Bertotti, G., Boro, H.: Multiscale fracture network characterization and impact on flow: A case study on the latemar carbonate platform. Journal of Geophysical Research: Solid Earth **120**(12), 8197–8222 (2015) https://doi.org/10.1002/2015JB011879

[37] Adler, P.M., Thovert, J.-F.: Fractures and Fracture Networks. Springer, Dordrecht (1999). https://doi.org/10.1007/978-94-017-1599-7

[38] Sahimi, M.: Flow and Transport in Porous Media and Fractured Rock, 2. edn. Wiley, Germany (2011). https://doi.org/10.1002/9783527636693

[39] Liu, R., Li, B., Jiang, Y., Huang, N.: Review: Mathematical expressions for estimating equivalent permeability of rock fracture networks. Hydrogeology Journal **24**(7), 1623–1649 (2016) https://doi.org/10.1007/s10040-016-1441-8

[40] Hadgu, T., Karra, S., Kalinina, E., Makedonska, N., Hyman, J.D., Klise, K.,

Viswanathan, H.S., Wang, Y.: A comparative study of discrete fracture network and equivalent continuum models for simulating flow and transport in the far field of a hypothetical nuclear waste repository in crystalline host rock. Journal of Hydrology **553**, 59–70 (2017) https://doi.org/10.1016/j.jhydrol.2017.07.046

[41] Kottwitz, M.O., Popov, A.A., Abe, S., Kaus, B.J.P.: Investigating the effects of intersection flow localization in equivalent-continuum-based upscaling of flow in discrete fracture networks. Solid Earth **12**(10), 2235–2254 (2021) https://doi.org/10.5194/se-12-2235-2021

[42] Müller, S., Schüler, L.: GSTools. https://github.com/GeoStat-Framework/GSTools (2019)

[43] Sandve, T.H., Berre, I., Nordbotten, J.M.: An efficient multi-point flux approximation method for Discrete Fracture–Matrix simulations. Journal of Computational Physics **231**(9), 3784–3800 (2012) https://doi.org/10.1016/j.jcp.2012.01.023

[44] Berrone, S., Pieraccini, S., Scialò, S.: On Simulations of Discrete Fracture Network Flows with an Optimization-Based Extended Finite Element Method. SIAM Journal on Scientific Computing **35**(2), 908–935 (2013) https://doi.org/10.1137/120882883

[45] Březina, J., Stebel, J.: Analysis of model error for a continuum-fracture model of porous media flow. In: Kozubek, T., Blaheta, R., Šístek, J., Rozložník, M., Čermák, M. (eds.) High Performance Computing in Science and Engineering, pp. 152–160. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40361-8_11

[46] Březina, J., Stebel, J., Exner, P., Hybš, J.: Flow123d. http://flow123d.github.com, repository: http://github.com/flow123d/flow123d (2011–2023)

[47] Durlofsky, L.J.: Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. Water Resources Research **27**(5), 699–708 (1991) https://doi.org/10.1029/91WR00107

[48] Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M., Farhan, L.: Review of deep learning. Journal of Big Data **8**(1) (2021) https://doi.org/10.1186/s40537-021-00444-8

[49] Zhang, A., Lipton, Z.C., Li, M., Smola, A.J.: Dive Into Deep Learning, (2020). https://d2l.ai

[50] Bednar, J.A., Crist, J., Cottam, J., Wang, P.: Datashader. https://datashader.org/ (2022)

[51] He, C., Yao, C., Jin, Y.-z., Jiang, Q.-h., Zhou, C.-b.: Effective permeability of fractured porous media with fracture density near the percolation threshold. Applied Mathematical Modelling **117**, 592–608 (2023) https://doi.org/10.1016/j.apm.2023.01.010

[52] Zhu, C., Wang, J., Sang, S., Liang, W.: A multiscale neural network model for the prediction on the equivalent permeability of discrete fracture network. Journal of Petroleum Science and Engineering **220**, 111186 (2023) https://doi.org/10.1016/j.petrol.2022.111186