AGU100 ADVANCING EARTH AND SPACE SCIENCE

Water Resources Research

RESEARCH ARTICLE
10.1002/2018WR022886

Key Points:
- We propose an algorithm of downscaling followed by segmentation to reconstruct the unresolved pore-space from XCT images
- The method is applicable to moderately unresolved, spatially heterogeneous but chemically homogeneous granular media
- The method is not based on the definition of arbitrary thresholds

Correspondence to:
I. Battiato,
ibattiat@stanford.edu

Citation:
Korneev, S. V., Yang, X., Zachara, J. M., Scheibe, T. D., & Battiato, I. (2018). Downscaling-based segmentation for unresolved images of highly heterogeneous granular porous samples. *Water Resources Research*, *54*. https://doi.org/10.1002/2018WR022886

Received 7 MAR 2018
Accepted 11 MAR 2018
Accepted article online 24 MAR 2018

Downscaling-Based Segmentation for Unresolved Images of Highly Heterogeneous Granular Porous Samples

S. V. Korneev[1], X. Yang[2], J. M. Zachara[3], T. D. Scheibe[4,5], and I. Battiato[1]

[1]Energy Resources Engineering, Stanford University, Stanford, CA, USA, [2]Geographical Science, Beijing Normal University, Beijing, China, [3]Fundamental and Computational Sciences Directorate, Pacific Northwest National Laboratory, Richland, WA, USA, [4]Energy and Environment Directorate, Pacific Northwest National Laboratory, Richland, WA, USA, [5]Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, WA, USA

Abstract Numerical simulations of pore-scale flow and transport in natural sediments require the knowledge of pore-space topology. Limited resolution of X-ray tomography is often insufficient to fully characterize pore-space structure within fine-grained regions. Single and multilevel threshold-based segmentation approaches are customarily employed to identify solid, pore and porous-solid regions by means of grey intensity thresholds. While the choice of cutoff thresholds is often arbitrary, it dramatically affects the effective properties and the dynamical response of the reconstructed porous structure. We propose an algorithm of downscaling, i.e., the process of increasing image resolution, followed by segmentation, i.e., the identification of different phases, to reconstruct the unresolved pore-space from XCT images of natural geological porous media. The method, applicable to moderately unresolved, chemically homogeneous granular media, is based on a map between local pixel porosity and pore size that does not rely on the definition of arbitrary thresholds and it allows to generate a high-resolution binary image of the porous medium from poorly resolved grey-scale images. First, we validate the method on synthetic unresolved images and compare their known pore-space distribution with the extracted one. Then, we consider a synthetic porous medium and compare the pore size distribution, conductivity, and tortuosity between the original and the reconstructed structures. Finally, we apply the method to extract the pore-space distribution from unresolved XCT images of two natural sediment columns and use it (i) to parametrize a capillary-bundle model and (ii) to estimate the hydraulic conductivity by matching breakthrough behavior of passive solute transport.

1. Introduction

Accurate reconstruction of the 3-D geometry is critical to model single (Ling et al., 2016; Scheibe et al., 2015a) and multiphase (Ling et al., 2017) pore-scale fluid flow and mass transport through highly heterogeneous natural porous media. While only a limited number of laboratory prototypes of multiscale imaging techniques exist (Andrä et al., 2013a, 2013b; Prodanović et al., 2014; Sok et al., 2010; Wildenschild & Sheppard, 2013), it is important to fully characterize the geometrical information across scales in an integrated fashion since flow and transport processes are often coupled across length scales that span different orders of magnitude (Arunachalam et al., 2015; Battiato, 2014; Battiato et al., 2009; Battiato & Tartakovsky, 2011; Boso & Battiato, 2013; Korneev & Battiato, 2016; Yousefzadeh & Battiato, 2017). One of the main difficulties lies in the uncertainty produced during the imaging experiment. The sources of uncertainty may include photon shot-noise, distortion of the optical system, low spatial resolution (Scheibe et al., 2015b; Soulaine et al., 2016), low sensitivity, and others (Boas & Fleischmann, 2012). These sources, when combined together, can significantly decrease image quality. In particular, subresolution porosity can lower the accuracy of flow and transport model predictions, regardless of the accuracy of the numerical solver (Boas & Fleischmann, 2012; Scheibe et al., 2015b; Soulaine et al., 2016). Prohibitive costs associated with techniques designed to improve image quality, such as adaptive optics (Davies & Kasper, 2012), sophisticated detectors and emitters (Wildenschild & Sheppard, 2013), and multiscale imaging techniques (Sok et al., 2010), limit their deployment in routine industrial, geosciences, and hydrological applications.

An alternative and less expensive approach is to mathematically postprocess the imaging data set. Recently, methods of noise reduction using filters (such as Gauss filter, nonlinear diffusion filter, and shock filter) have been successfully applied to digital images (Kaestner et al., 2008). Nevertheless, the problem of unresolved porosity, known as the partial volume effect (PVE), is more difficult to resolve. PVE occurs when the smallest pore spaces of a sample are below the device resolution and the information about subvoxel variations in chemical composition and/or phase is averaged over a single voxel. If the PVE voxel includes both solid and void phases it is usually referred to as a porous-solid phase (Scheibe et al., 2015b). The image intensity of a voxel with porous-solid phase contains information about subresolution porosity and pore-space geometry that cannot be explicitly resolved.

Reconstruction of a digital porous structure where direct numerical simulations can be performed or physical multiscale models parametrized requires a preliminary segmentation step, i.e., the identification of different phases (pore, solid, porous solid, etc.) from the original image. Binary segmentation approaches neglect the presence of the porous-solid phase and reconstruct the synthetic pore structure by imposing a threshold on the intensity histogram to separate solid phase from pore space (Lindquist et al., 1996). Numerous studies have demonstrated the dramatic impact that such an approximation has on characterization of the overall conductivity as well as the dynamical response of natural porous media, e.g., Soulaine et al. (2016). To overcome this challenge, ternary segmentation approaches have been employed where two intensity threshold values have to be determined to distinguish among solid, porous-solid, and void phases. Yet the choice of thresholds can be ill defined as the boundaries between phases may not be clearly detectable on the image and the intensity histogram (Iassonov et al., 2009; Scheibe et al., 2015b; Soulaine et al., 2016; Wildenschild & Sheppard, 2013). Scheibe et al. (2015b) show how breakthrough curves of passive transport in a reconstructed geologic porous medium are greatly affected by the type of segmentation and threshold values employed. Once the three phases have been identified, the porous-solid phase is generally assumed spatially homogenous and its conductivity estimated through empirical relationships, such as the Carman-Kozeny equation, that are not universal and may not work for a highly heterogeneous sample (Iassonov et al., 2009; Scheibe et al., 2015b; Soulaine et al., 2016; Wildenschild & Sheppard, 2013). Yet, an accurate knowledge of the pore geometry is essential to calculate hydraulic conductivity, hydrodynamic dispersion, or effective surface reaction rates (Mehmani & Tchelepi, 2017; Yang et al., 2016; Yoon et al., 2015) in at-scale and multiscale models (Battiato, 2016; Battiato et al., 2011; Sun et al., 2012; Yousefzadeh & Battiato, 2017). The geometry of the porous-solid phase can also be recovered by stochastic reconstruction (Gerke et al., 2015). This approach does not require a threshold value or empirical parameters, but it is computationally intensive since it is based on the solution of a multidimensional minimization problem (Mohebi et al., 2009; Mohebi & Fieguth, 2008; Wildenschild & Sheppard, 2013). While relatively new to geosciences, the problem of subresolution reconstruction remains a challenging step in tomographic image processing and has long been a goal in the analysis of medical images such as brain MRI (Bultreys et al., 2014; Hwang & Felix, 2002; Wildenschild & Sheppard, 2013, and references therein). More detailed reviews of modern image processing algorithms in geosciences and hydrology can be found in Bultreys et al. (2016), Wildenschild and Sheppard (2013) and Kaestner et al. (2008).

Here we propose and validate an algorithm of downscaling followed by segmentation to extract information about pore geometry from unresolved grey-scale images of a highly heterogeneous granular sample, which we will refer to as downscaling-based segmentation (section 2). The downscaling step consists in increasing the image resolution, while the segmentation step involves the identification of different phases, e.g., solid and void. In the absence of a porous-solid phase, the segmentation step corresponds to binarizing the image. The central element of the approach is a continuous mapping of the low-resolution pixel intensity (section 2.1) into the pore space of a high-resolution binary image (section 2.2), followed by subsequent segmentation of the resulting pore space into interconnected bounded regions. The latter step allows one to extract relevant pore-scale statistics and spatial distribution of pores. Unlike the commonly used threshold-based segmentation which ignores subresolution information (Iassonov et al., 2009), the mapping applied here retains more than one bit information from a voxel by means of postulating a morphological relationship between the local pixel porosity, i.e., the average volume of the pore in any given pixel, and the characteristic size of the pore space at that location. The algorithm has a simple formulation and high computational efficiency for moderately unresolved images and chemically homogeneous well-packed granular samples. We first test the method using a poorly resolved synthetic image with known pore size distribution. To obtain a low-resolution grey-scale image, we generate a synthetic high-resolution binary image of a heterogeneous porous medium and upscale it using volume averaging. We also add Gaussian noise to the low-resolution image. In section 3, we apply the downscaling algorithm followed by segmentation to extract the reconstructed pore-space distribution. We also compare the extracted pore-space distribution from the downscaling-based and the threshold-based segmentations with the original distribution. The comparison shows that the downscaling-based approach is more accurate than the threshold-based method. Finally, we consider two examples where we apply the model to calculate effective parameters in two three-dimensional media. In the first example, we determine the effective conductivity of a synthetically generated porous medium and compare both its pore-scale statistics, tortuosity and conductivity with the reconstructed one (section 3.1). In the second example (section 3.2), we demonstrate how the subresolution information extracted by the algorithm (e.g., pore size distribution) can be employed to parametrize upscaled flow and transport models. Specifically, we apply the downscaling-based segmentation approach to poorly resolved XCT images of two natural geological samples (Zachara et al., 2013). By utilizing the extracted geometry to parametrize a capillary-bundle flow model, we are able estimate the hydraulic conductivity and solute transport behavior of the samples. The comparison with experimental results shows that the method better models the unresolved porosity and its impact on macroscopic behavior than a segmentation-based approach. We provide a summary of the main results in section 4.

2. Algorithm

XCT scanning is a widely used technique for nondestructive and noninvasive three-dimensional (3-D) imaging of geological samples (Bultreys et al., 2016; Wildenschild & Sheppard, 2013). The output of the XCT experiment provides a 3-D distribution of the relative attenuation of the scattered X-ray beam, where the intensity depends on the distribution of various phases (solid or void) and mineral composition across the sample. Here we focus on chemically homogeneous samples. Generalizations of the approach to chemically heterogeneous samples are subject of current investigations.

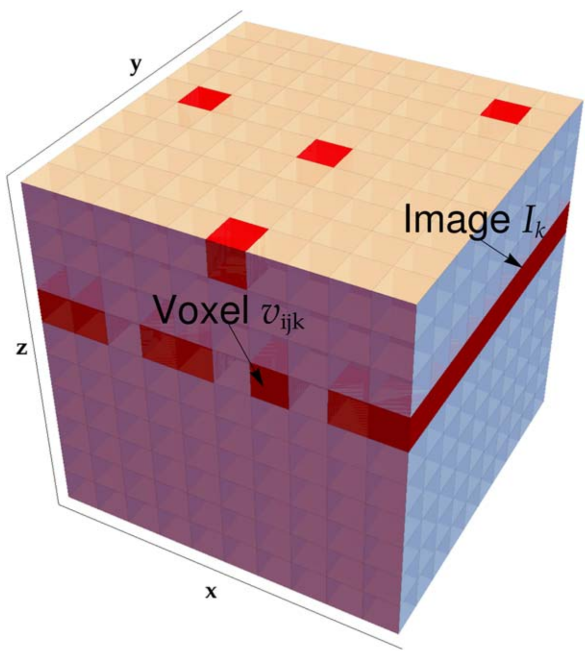


Figure 1. Representation of a tomographic 3-D image S, where the voxel is identified by an array of three indeces (i, j, k), representing the voxel coordinates in the x, y, and z direction, respectively, and its intensity is v_{ijk}. The image S can be represented as a collection of 2-D images I_k in the xy-plane.

2.1. Step 1: Intensity-Porosity Mapping

We define the result of an XCT measurement as a stack $\mathcal{S}$ of 2-D images $\mathcal{I}_k$, $\mathcal{S}=\{\mathcal{I}_k,\ k=1,\ldots,N_k\}$, where $\mathcal{I}_k=\{(i,j,k),\ i=1,\ldots,N_i,\ j=1,\ldots,N_j\}$ is a tagged image file format (TIFF), N_k is the number of images, $N_i \times N_j$ the image size and $v_{ijk} \in [0,1]$ is voxel (i, j, k)'s intensity (see Figure 1). The rescaled voxel size is $1 \times 1 \times \Delta z$, where Δz is the distance between two contiguous 2-D images $\mathcal{I}_k$ and $\mathcal{I}_{k+1}$, and the corresponding 2-D pixel size in the xy-plane is 1×1. Since tomographic images are obtained slice wise (Lindquist et al., 1996), the method is performed on individual slices. We process stacks of 2-D images, and we will refer to 2-D "pixels" rather than 3-D voxels. Pixel intensity V is a Gaussian random process in the limit of large photon count (Lei, 2012), i.e.,

$$f_V(v)=a_s f_G(v;v_s,\sigma_s^2)+\sum_{p=1}^{N_p} a_p f_G(v;v_p,\sigma_p^2)+a_v f_G(v;v_v,\sigma_v^2), \qquad (1)$$

where $f_V(v):=\Pr(V=v)$ and $f_G(v;\mu_i,\sigma_i^2)=(2\sigma_i^2\pi)^{-1/2}e^{-(v-\mu_i)^2/2\sigma_i^2}$ is a Gaussian distribution with mean μ_i and variance σ_i^2, $i=\{s,v,p\}$.

The first step of the algorithm consists in mapping the image grey scale into local (pixel) porosity values. If the signal-to-noise ratio is high enough, the solid and void phases can be identified as maxima of the data intensity histogram: this is achieved by fitting two Gaussian distributions to the left and right maxima of the intensity histogram to capture the mean grey intensity v_s and v_v associated with the

10.1002/2018WR022886

solid and pore phases. Otherwise, linear regression between $(v_v - v_s)$ and v_s or deconvolution of the histogram (Fister et al., 2007) can be employed to determine v_v and v_s. Once v_v and v_s are determined, the pixel (i, j)'s porosity ϕ_{ij} of the original image $\mathcal{I}$ can be determined through the linear map,

$$\phi_{ij}(v_{ij}) = \frac{v_{ij} - v_s}{v_v - v_s}, \tag{2}$$

similar to the definition of Hounsfield units, and other linear mappings (Sok et al., 2010). Equation (2) ensures that $\phi(v_s)=0$ in the solid phase and $\phi(v_v)=1$ in the void phase. The porosity distribution is Gaussian, i.e.,

$$f_\Phi(\phi) = a_s f_G\left(\phi; 0, \hat{\sigma}_s^2\right) + \sum_{p=1}^{N_p} a_p f_G\left(\phi; \phi_p, \hat{\sigma}_p^2\right) + a_v f_G\left(\phi; 1, \hat{\sigma}_v^2\right), \tag{3}$$

and its expected value is

$$\bar{\phi} = \sum_{p=1}^{N_p} a_p \phi_p + a_v = \frac{1}{N_i N_j} \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \phi_{ij}. \tag{4}$$

This approach provides the porosity distribution of each image, does not require an arbitrary threshold for phases' identification, and allows one to attribute to the porous-solid phase spatially heterogeneous properties (i.e., porosity) depending on the pixel local intensity. We emphasize that equation (4) provides a fit-free estimate of the porosity that can be calculated directly from the image and compared with the experimental value for validation.

2.1.1. Synthetic Image Generation

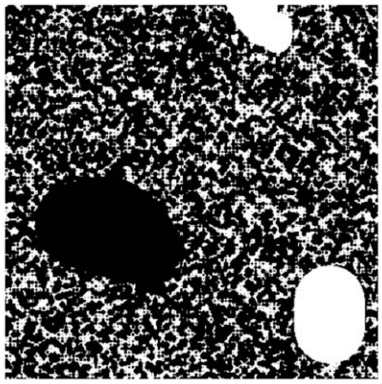
We test the first step of the method on a synthetically generated low-resolution image (Figure 2b) with known underlying pore-space distribution (Figure 2a). The low-resolution grey-scale image of size 400 $\times$ 400 pixels (Figure 2b) is obtained by adding a Gaussian noise with zero mean and 0.035 standard deviation to the spatially upscaled synthetic high-resolution binary image 4000 $\times$ 4000 pixels in size (Figure 2a). The noise standard deviation was chosen such that the final pixel intensity histogram of the low-resolution image in Figure 2c would appear as realistic as possible (see e.g., histogram of a real XCT image in Figure 18). The total porosity of the high-resolution image is $\phi=0.253$. In order to ensure that the low-resolution image contains unresolved porosity, its resolution (i.e., its pixel size) should be coarser than the average pore width $\bar{a}$ of the original (high-resolution) synthetic image. The average pore size can be determined through segmentation of the image in Figure 2a. We employ a segmentation approach known as connected-component labeling. It can be performed utilizing fast and efficient algorithms (Samet & Tamminen, 1988). In such a method, the void space of the binary image is segmented in interconnected regions bounded by the solid phase, where each connected region is here referred to as *pore*. An interconnected region is defined as a two-dimensional area (or three-dimensional volume) such that any arbitrary pair of points within the region can always be connected by a continuous curve. Even though such a segmentation significantly overestimates effective pore size in loosely packed porous media where the void spaces may merge into large connected pores (see Figure 3a), the approach provides a good approximation of pore-space distribution for well-packed granular (geologic) porous media (see Figure 3b). We emphasize that using this segmentation approach, a *pore* may be a nonconvex void space (i.e., for every pair of points within a pore, every point on the straight line segment that joins the pair of points may not be within a void space). One such example is shown in Figure 3b, where one of the pores, i.e., the interconnected region in the top-right corner, contains one solid grain in it.

After segmentation (Figure 3c), the average pore width $\bar{a}$ of the high-resolution image is determined as $\bar{a}=\bar{a}_p=4.0$, the arithmetic average of the pores' characteristic width $a_p=\sqrt{S_p}$, where S_p is the pore area and the subscript p is the pore index. The resolution of the low-resolution image is set to 10, 2.5 times coarser than the average pore width $\bar{a}$ of the original synthetic image.

2.1.2. Gaussian Fit and Intensity-Porosity Map

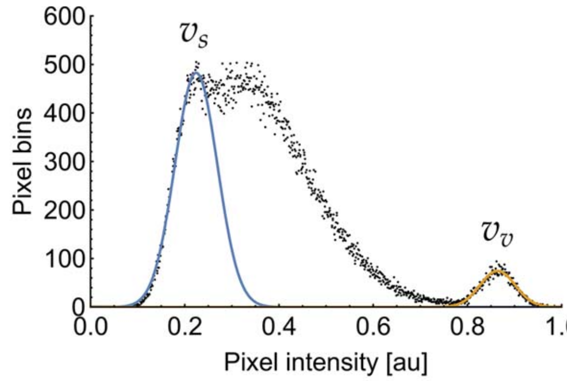
In Figure 2c, we show the intensity histogram of the synthetically generated low-resolution image with known underlying pore-space distribution. The intensities associated with the solid and void phases v_s and v_v, respectively, are determined by fitting two Gaussians to the left and right maxima of the intensity

10.1002/2018WR022886



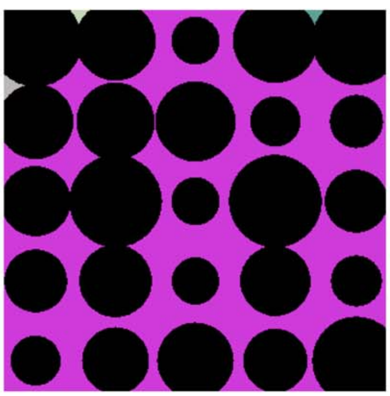
(a) High-resolution synthetic image

(b) Low-resolution synthetic image



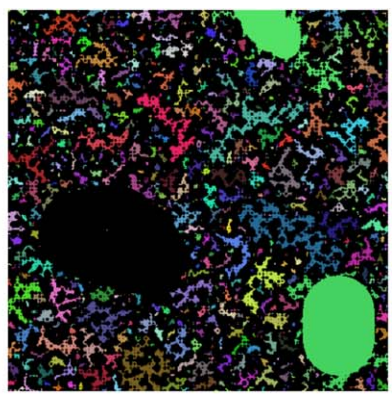
(c) Pixel intensity histogram in arbitrary units and Gaussian fits

Figure 2. (a) Original high-resolution synthetic binary image (4,000 × 4,000 pixels). (b) Upscaled low-resolution image with Gaussian noise (400 × 400 pixels). (c) Intensity histogram of image in Figure 2b (black dots) with Gaussian fits (solid lines) to identify the grey scale intensities v_s and v_v associated with the solid (solid blue line) and void phases (solid yellow line), respectively.



(a) Loosely-packed porous medium

(b) Well-packed porous medium



(c) High-resolution synthetic image

Figure 3. Connected-component labeling segmentation for a (a) loosely and (b) closely packed synthetic porous medium, and (c) the high-resolution synthetic image of Figure 2a. In all figures, the black corresponds to the solid phase. The "pores" determined through connected-component labeling segmentation are represented in different colors. Such a segmentation algorithm overestimates the characteristic pore size in loosely packed porous media (Figure 3a), while it provides an accurate estimate of the pore size distribution for closely packed porous media (Figure 3b). (c) Segmentation result on the synthetic image of Figure 2a. Different colors identify different pores p (or connected nonconvex void regions) of area S_p and characteristic size $a_p = \sqrt{S_p}$.

10.1002/2018WR022886

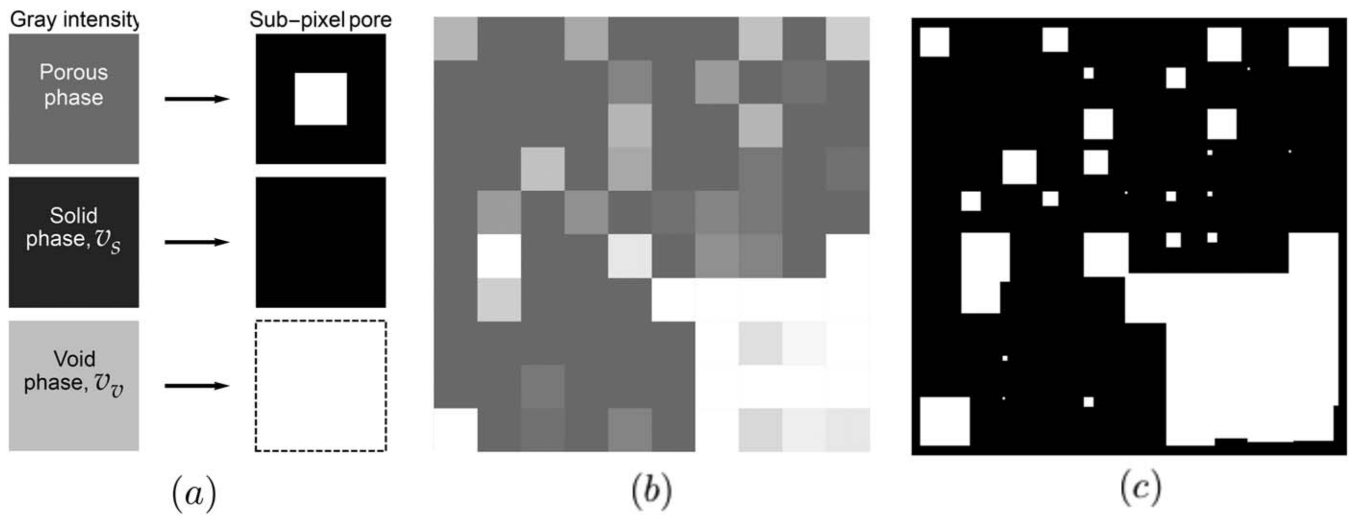


Figure 4. (a) Schematic representation of the downscaling step where (left) the pixel porosity ϕ_{ij} is mapped into (right) a square pore of size $w_{ij} \times w_{ij}$, i.e., a binary distribution of voids (in white) and solid (in black) phases at that given pixel. Pixels with porosity equal to 1 and 0 correspond to an entirely void (white) or solid (black) pixel. (b, c) Example of downscaling from (b) a pixel porosity distribution to (c) a binary (black and white) image, where void pixels may overlap and merge to form larger connected regions.

histogram as shown in Figure 2c. The fitted intensity values are $v_s = 0.22$ and $v_v = 0.864$. The local pixel porosity of the low-resolution image is then determined using the intensity-porosity relation (2). The estimated porosity using equation (2) is $\bar{\phi}=0.225$ and compares well with the porosity $\phi=0.253$ of the original image.

2.2. Step 2: Downscaling-Based Segmentation

Once the pixel porosity of the image is estimated, the spatial distribution of void and solid phases must be reconstructed in a binary image such that direct numerical simulations can be run at the pore-scale. This requires a downscaling step to map local pixel porosity onto a binary higher-resolution image. Since the underlying pore geometry of the solid porous phase is unknown, a model is necessary to map local pixel porosity onto an equivalent (binary) distribution of void and solid phases at any given pixel. In the following, we propose two different models used to reconstruct the pore structure (i.e., the spatial distribution of voids and solids) within each pixel. In both models, we assume that each pixel contains a single pore whose size depends only on the pixel intensity as described in Figure 4. This assumption is deemed reasonable for chemically homogeneous granular media and a moderately unresolved image. While a downscaling step cannot capture the full topological information of the subpixel pore structure, the downscaling step can be designed to extract at least part of such information (e.g., the characteristic pore size of subresolution porosity).

Beyond running direct numerical simulations, the downscaled high-resolution image can be employed (i) to parametrize transport models as shown in section 3, (ii) to estimate the porous medium conductivity, including that associated with the porous-solid phase, without relying on empirical relationships, e.g., Carman-Kozeny equation and/or (iii) to perform numerical upscaling after the porosity map (Step 1 of the algorithm), (iv) to analyze the pore-space topology and estimate, e.g., pore connectivity (Bernabe et al., 2010; Lindquist et al., 1996).

2.2.1. Physical Subpixel Pore Model

We assume that any arbitrary pixel (i, j) (of dimensionless area 1×1) contains a square pore of area w_{ij}^2 located at the center of the pixel (see Figure 4a). Since the pixel intensity v_{ij} depends linearly on the mass density and, consequently, on the pore area w_{ij}^2, the following map, here referred to as physical model, can be defined:

$$\begin{cases} w_{ij}(\phi_{ij}, \delta)=\sqrt{(1+\delta)\phi_{ij}}, & \phi_{ij} > 0, \\ w_{ij}(\phi_{ij})=0, & \phi_{ij} \le 0. \end{cases} \tag{5}$$

Equation (5) guarantees that $w_{ij} = 0$ when $v_{ij}=v_s$, and $w_{ij}=1+\delta$ when $v_{ij}=v_v$, where the rescaled pixel width is 1, and δ is a fitting parameter, determined by matching the total porosity $\bar{\phi}=\langle\phi_{ij}\rangle$ of the original image

10.1002/2018WR022886

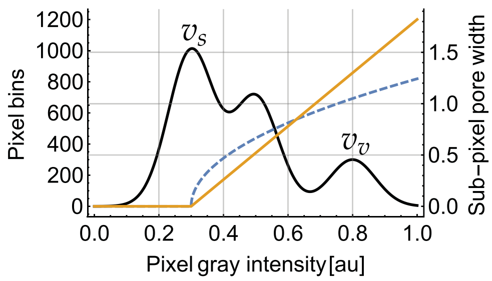


Figure 5. The solid black curve shows a pixel intensity histogram, where the left maximum v_s corresponds to the solid phase and the right maximum v_v to the void phase. The dashed blue curve shows the physical subpixel pore model (see equation (5)), and the solid yellow line shows the morphological model (see equation (6)).

(as calculated from Step 1) with the total porosity $\bar{\phi}^*$ of the binary image. The parameter δ should ensure that neighboring subpixel pores form a connected region when the binary image is segmented. This is critical to preserve connectivity of large pores after the downscaling step. Yet, it is easy to observe that, in the physical model, $\bar{\phi}^*=(1+\delta)\langle\phi_{ij}\rangle=(1+\delta)\bar{\phi}$, i.e., $\delta\approx 0$. As a result, neighboring subpixel pores may not merge into interconnected bounded void spaces in presence of noise and the downscaled image may lose connectivity. To overcome this problem, a morphological model is introduced.

2.2.2. Morphological Subpixel Pore Model

The aim of the morphological model is to smooth the random intensity variations by overestimating the subpixel pore width inside the void spaces where porosity is close to 1 (see Figure 5). Under the assumption that the subpixel pore width w_{ij} linearly depends on the pixel porosity ϕ_{ij}, the morphological subpixel model can be expressed as

$$\begin{cases} w_{ij}(\phi_{ij},\,\delta)=(1+\delta)\phi_{ij}, & \phi_{ij}>0, \\ w_{ij}(\phi_{ij})=0, & \phi_{ij}\le 0, \end{cases} \tag{6}$$

where the pore width is 0 for the solid phase v_s, and $1+\delta$ for the void phase v_v. This corresponds to a morphological dilation (Aoud, 2014) of the void space compared to the physical model. While the standard dilation algorithm smoothes small scale variations of the geometry and thus preserves the connectivity, it does not preserve the porosity. In the morphological subpixel model the fitting parameter δ is, again, determined by matching $\bar{\phi}=\langle\phi_{ij}\rangle$, the total porosity of the original image with the total porosity of the binary image $\bar{\phi}^*(\delta)=(1+\delta)^2\langle\phi_{ij}^2\rangle$, i.e., $\delta=\sqrt{\langle\phi_{ij}\rangle/\langle\phi_{ij}^2\rangle}-1\neq 0$.

2.2.2.1. Validation on a Synthetic Image

We test the downscaling by means of the physical, Figure 6a, and morphological model, Figure 6b, on the synthetic image in Figure 2b (with underlying known pore structure provided in Figure 6a) where white and black pixels correspond to voids and solids, respectively, and grey to a solid porous phase with $0<\phi<1$. After downscaling, we segment the generated binary images to highlight the connected pore space. We emphasize that the physical and morphological models map a real number (porosity) into a real number (pore width). Once the pore width associated with each pixel is known, the 2-D array of pore widths (real numbers) is discretized to obtain a 2-D binary array (binary numbers). The increase in resolution originates during this discretization step. In this example, and elsewhere, we use a relatively high resolution to

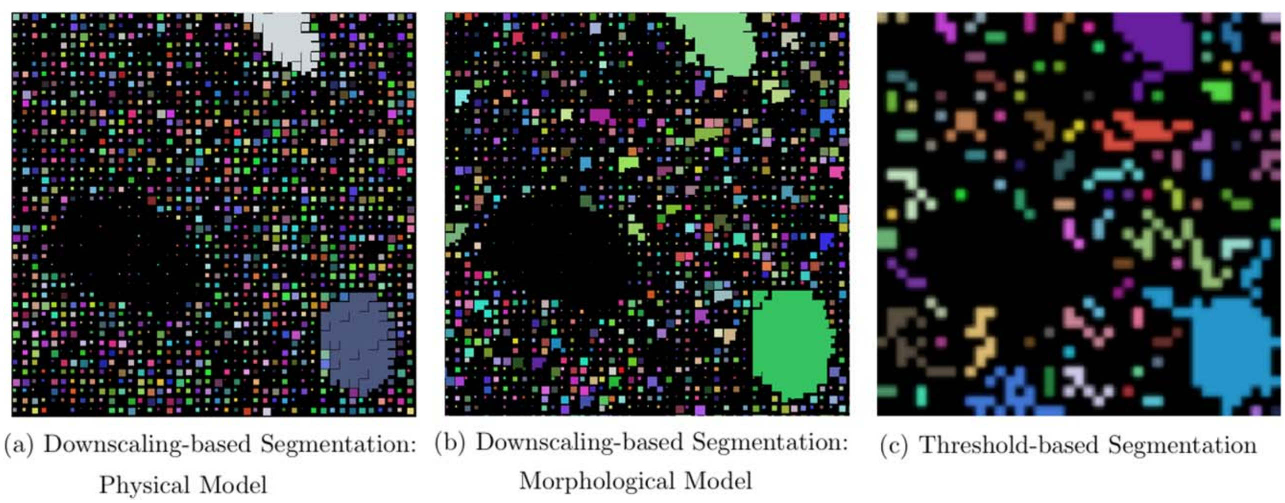
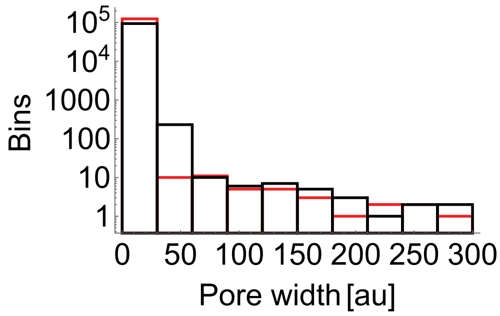
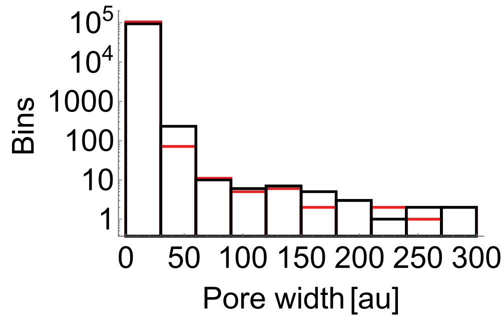


Figure 6. Segmentation results from a (a) downscaling-based physical model, (b) downscaling-based morphological model, and (c) threshold-based method, Different colors represent individual pores.

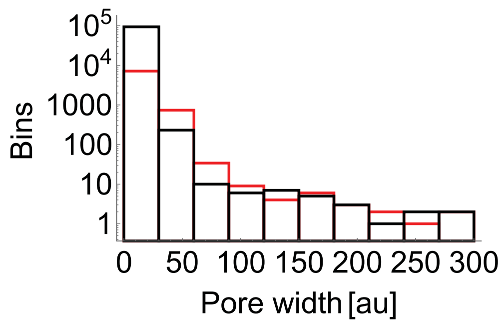
10.1002/2018WR022886



(a) Downscaling-based segmentation: Physical model



(b) Downscaling-based segmentation: Morphological model



(c) Threshold-based segmentation

Figure 7. Histograms of pore width extracted after segmenting the low-resolution image 2(b) with different methods (red bins): (a) downscaling-based segmentation through the physical model, (b) downscaling-based segmentation through the morphological model, and (c) threshold-based segmentation.

be able to explicitly resolve boundaries between neighboring subpixel pores. Then, we calculate the pore-space statistics, i.e., the distribution of the pore-space characteristic size and its average. For comparison, we also apply a standard threshold-based segmentation approach (Figure 6c) where the threshold value is selected such that the total porosity of the resulting binary image matches the porosity of the low-resolution image. The average pore size after segmentation is $\bar{a}=4.5$ pixel, $\bar{a}=4.3$ pixel, and $\bar{a}=17.5$ pixel for the physical, the morphological, and the threshold-based segmentation approaches, respectively. The morphological model provides the best estimate of the true pore size, $\bar{a}=4$. The results of the three approaches are shown in Figure 6. In Figure 7, we compare the pore size histogram obtained from segmenting the original high-resolution synthetic binary image (solid black histogram) with the pore size histograms of the binary images obtained from a (i) downscaling-based segmentation with the physical and morphological models (Figures 7a and 7b, respectively) and (ii) a threshold-based segmentation (Figure 7c). The histogram comparison reveals that downscaling using the morphological model (6) gives the best approximation of the original histogram. The physical model (5) underestimates the number of pores of intermediate size, e.g., pores with width between 50 (au) and 100 (au), by a factor of 10. Threshold-based segmentation underestimates the number of the smallest pores by factor of 10. Once the statistics of the pore space are extracted, they can be used to parametrize flow and transport models as demonstrated in the following section.

Finally, we test how both models preserve connectivity of large pores (i.e., pores above the resolution threshold). We generate a synthetic high-resolution binary image with a single large pore embedded in a fine-grained porous matrix (see Figure 8a). The binary image is first upscaled using volume averaging (see Figure 8b). We then add a Gaussian random variable with zero mean to the upscaled (low-resolution grey scale) image to model noise (see Figure 8c). We use four different values of the standard deviation ranging from 0.02 to 0.1. We scale the intensity of the low-resolution image such that the maximum corresponding to the void phase (white) is located at 1 and that of the solid phase (black) at 0. Finally, we extract the large pore from the low-resolution image using the downscaling method with physical and morphological models (see example in Figure 9). For each set of parameters, we run the simulations 1,000 times and count the number of successful extractions. The results are listed in Table 1. Our results show that the morphological model always leads to 100% successful extractions of the large pore, while the extraction probability for the physical model depends on the noise standard deviation σ (specifically, extraction probability increases as σ decreases, as expected). This can be explained as follows: the morphological model, compared to the physical model, overestimates the size of the pores for values of porosity approaching 1, and underestimates the size of the pores when porosity approaches 0, as shown in Figure 4. This helps preserving intra-slice connectivity of large pores, while enhancing separation of neighboring small pores.

In the following section, we present two applications to 3-D porous media.

3. Example Applications

In the following, we present two examples where we use the downscaling-based segmentation in conjunction with (pore-scale and macroscale) flow (and transport) models to estimate effective properties of the

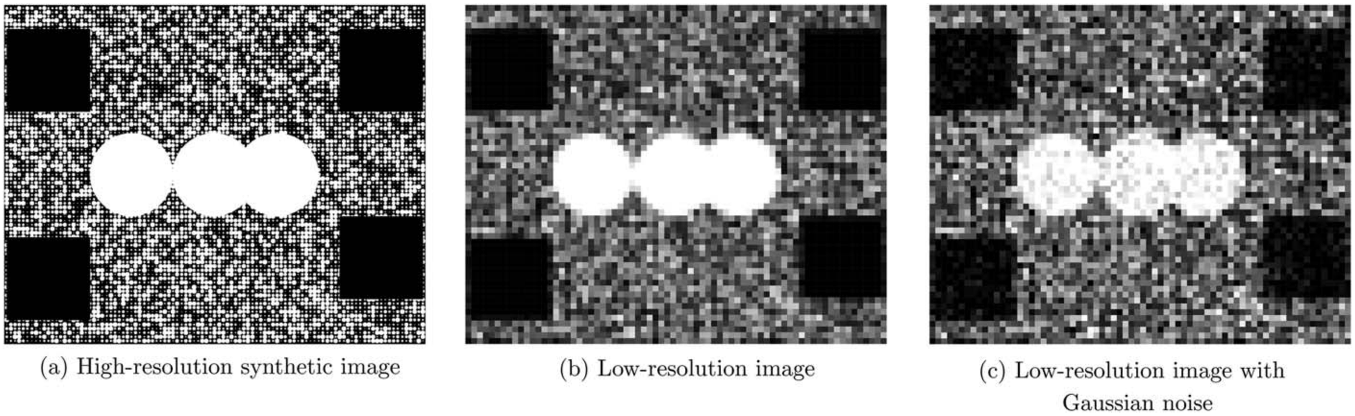


Figure 8. High-resolution synthetic binary image of size 720 × 580 pixels with a single large pore on the background consisting of small pores, where white is the void phase, black is the solid. (b) Low-resolution image of size 72 × 58 obtained by the volume averaging. (c) Upscaled image with the additive Gaussian noise which has with zero mean value and 0.1 the standard deviation.

reconstructed porous medium. In the first example (section 3.1), we consider a synthetically generated 3-D porous medium. After image processing, we reconstruct the pore-scale topology based on the downscaling approach, and then use pore-scale simulations on the original and reconstructed geometries to calculate the sample effective conductivity and tortuosity. In the second example (section 3.2), we show how the pore-scale distribution extracted after image processing from XCT scans of two realistic porous media can be used to parametrize an upscaled (capillary-bundle) model. We would like to emphasize that the purpose of these examples is to show how the information extracted by the downscaling-based segmentation can be used either to run direct numerical simulations on the reconstructed geometries (Example 1) and/or to improve the parametrization of simplified (upscaled) models, if direct numerical simulations are not an option due to high computational burden (Example 2).

3.1. Example 1: Synthetic Porous Medium

We consider a 3-D synthetic medium, constituted of random curved channels, with variable cross sections along the channel's length, connecting the top and the bottom of the domain. The 3-D geometry of a channel i is fully described by the coordinates of the channel's center $\{x_i(z), y_i(z)\}$ in the xy-plane and its radius $r_i(z)$ as

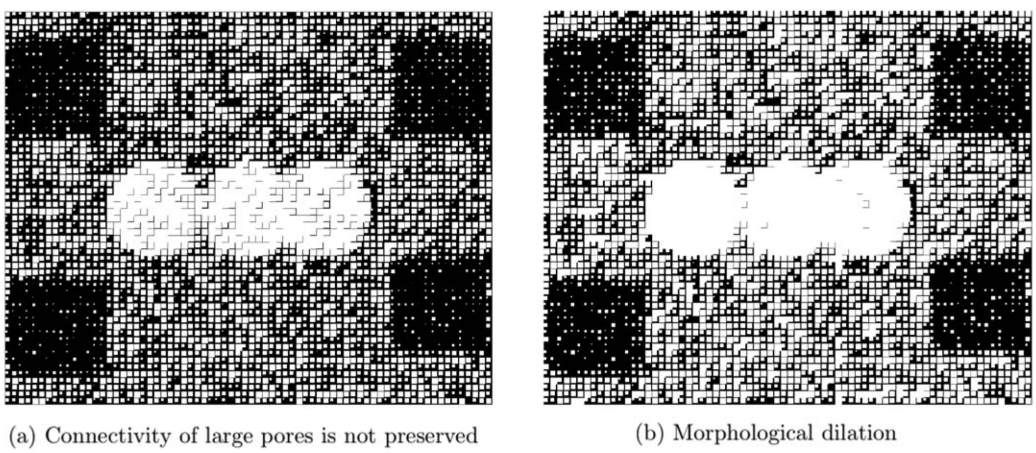


Figure 9. (a) Example of how the connectivity of the large pore is not preserved after the downscaling. (b) The morphological dilation restores the connectivity of the large pore, but changes the porosity. On both images, the black corresponds to the solid phase and white to the void.

Table 1
Probabilities of the Extraction of the Large Pore After Downscaling With the Physical and the Morphological Models for Different Values of White Noise Standard Deviation σ

Σ	Physical model	Morphological model
0.1	0.835	1
0.07	0.883	1
0.05	0.882	1
0.02	0.943	1

$$\{x_i(z), y_i(z)\}=f_i(z),$$
$$(x-x_i)^2+(y-y_i)^2 < r_i(z)^2,\ z \in [0{:}1], \tag{7}$$

where x, y, and z are the horizontal and the vertical coordinates, and $f_i(z)$ is a Bézier curve generated from a set of random points. Specifically, we generate 300 random Bézier curves for the horizontal coordinates $\{f_i(z),\ i=1,\ldots,300\}$ and 300 curves for the radius $\{r_i(z),\ i=1,\ldots,300\}$. Using the constrains (7) and the generated curves, we create 201 black and white images with dimensions 250 × 250 pixels. The generated set of 250×250×201 pixels represents a 3-D porous medium with vertical and horizontal heterogeneities. An example image is shown in Figure 10 and the volumetric render of the 3-D porous structure in Figure 11. The channels average diameter across the whole domain is 20 pixels. To synthetically model unresolved porosity, we reduce the resolution of each image by a factor of 10 using volume averaging. The dimensions of the upscaled porous medium are 25×25×210 pixels. We also add Gaussian noise with zero mean and standard deviation 0.01 to each upscaled image. Then, we reconstruct the low-resolution images using the downscaling algorithm and obtain a reconstructed porous medium whose dimensions are the same as those of the original domain. Figure 12 shows an example of the low-resolution and corresponding downscaled image for one cross section. The pore size distributions of the original and reconstructed domains are shown in Figure 13: despite the resolution of the image has been decreased by a factor of 10, the reconstruction algorithm can capture well the original pore size distribution.

In order to calculate conductivity and tortuosity, we perform a numerical Darcy's experiment using the original and the reconstructed porous domains. We use Cartesian grids while solving the dimensionless Stokes equation with an additional momentum loss term in the solid phase to ensure numerical stability, i.e.,

$$\nabla^2\mathbf{u}+\frac{\alpha}{R}\mathbf{u}=\nabla p,$$
$$\nabla\cdot\mathbf{u}=0, \tag{8}$$

where α is the phase indicator function, $\mathbf{u}$ is the flow velocity nondimensionalized by a characteristic flow velocity U, R is a momentum loss coefficient rescaled by L^2, with L the characteristic size of the flow domain, and p is the pressure rescaled by $L/(\mu U)$, with μ the fluid dynamic viscosity. In the void and solid phases, $\alpha=0$ and $\alpha=1$, respectively. Equations (8) are supported by inflow and outflow boundary conditions at the top and bottom of the computational domain. The dimensions of the computational domain are 2.5×2.5×2.01, and the indicator function α is set to either 1 or 0 in the black and white pixels, respectively. We set $R:=10^{-6}$, which corresponds to nearly impermeable solid boundaries while ensuring the stability of the numerical solver. We solve equations (8) using the finite volume framework OpenFoam. We modify the incompressible solver simpleFoam by adding the loss term in the momentum equation. After performing the numerical Darcy's experiment, we calculate the value of conductivity and tortuosity for the original and reconstructed domains, see Table 2. Tortuosity is obtained by measuring the normalized length of the streamlines. Figure 14 shows the streamlines in the original and reconstructed domains.

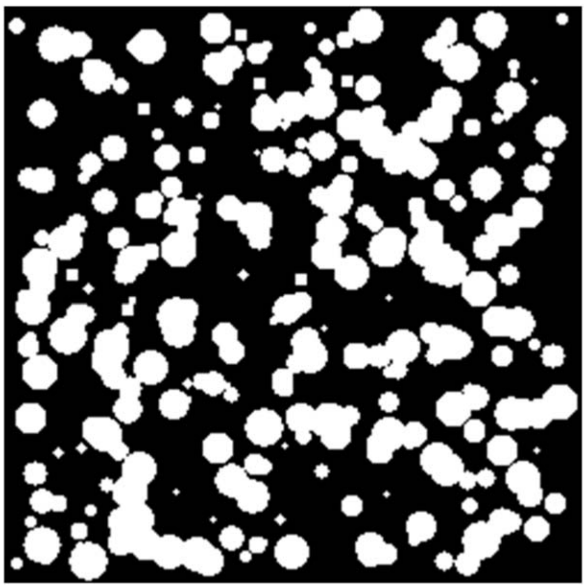


Figure 10. First image of the 3-D synthetic porous medium. The image dimensions are 250 × 250 pixels. White pixels corresponds void space and black pixels to the solid phase.

3.2. Example 2: Real Geologic Media

Direct numerical simulations of flow and transport (to calculate, e.g., effective conductivity and dispersion coefficient) may be computationally prohibitive, especially if the reconstructed core size is relatively large. Here we show that the pore-scale distribution extracted from the downscaling algorithm can be used to parametrize (reduced-order) macroscale models when the reconstructed domain

10.1002/2018WR022886

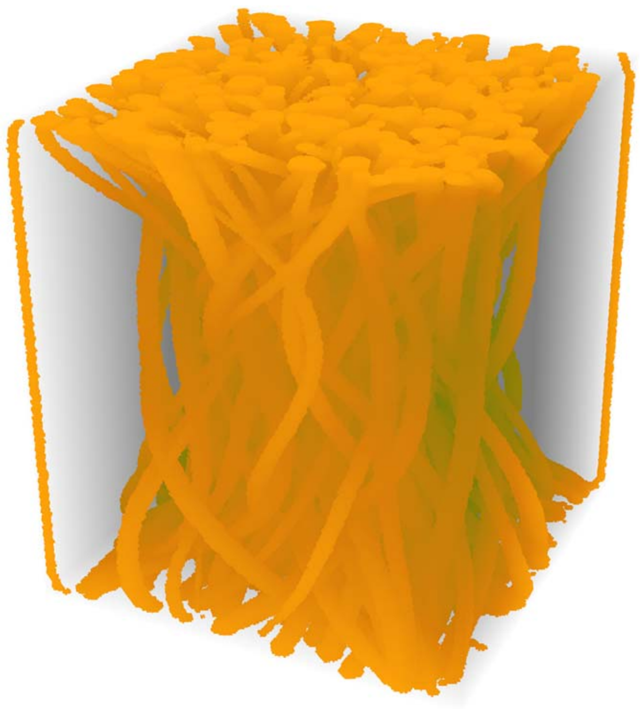


Figure 11. Volumetric render of the 201 synthetically images (not in scale). Dimensions of each image are 250 × 250 pixels.

is too large and render DNS prohibitive. Specifically, we apply the proposed image processing algorithm to XCT data from a natural porous medium, and use the extracted pore space to, first, fit the breakthrough curves of a passive tracer and, then, to predict the measured hydraulic conductivity. We employ a capillary-bundle model to parametrize hydraulic conductivity and passive transport features. The bundle's characteristic length scales are extracted either from the proposed image processing algorithm or from a threshold-based segmentation approach. The conductivities predicted in both cases are then compared to experimental measurements. We emphasize that the proposed framework is not restricted to the specific capillary-bundle model, chosen here for demonstration only, but can be used with any arbitrary macroscopic/reduced-order model of one's choice, where parameter fitting can be improved by a more accurate estimate of, e.g., the pore size distribution.

Two subsurface sediment columns C6197A and C6203A collected from boreholes in the 300 Area of the U. S. Department of Energy Hanford Site in South-Central Washington State (Scheibe et al., 2015b) are imaged by XCT scanning in the longitudinal direction to obtain $N_k = 1{,}000$ 2-D grey-scale TIFF images of the cross-sectional area of the cores. Each 2-D image has dimensions $N_i \times N_j = 1,024 \times 1,024$ pixels. The images of the cores C6197A and C6203A are stored as two stacks $\mathcal{S}_1$ and $\mathcal{S}_2$ of TIFF images, respectively. The volume render (for both columns) is shown in Figure 15 and a sample 2-D image in Figure 16. Flow and transport experiments were conducted on the columns (Scheibe et al., 2015b) to estimate porosity, hydraulic conductivity, and solute transport behavior (Table 3). Further details of the experimental setup can be found in Scheibe et al. (2015b). In Figure 17, we provide a schematic of the steps necessary to go from an XCT image to predictions of flow and transport processes. Each step is analyzed in detail in the following.

3.2.1. Step 1: Intensity-Porosity Mapping

Before processing the images, we crop the region around the rim of the column representing the container (circular ring in Figure 16a). We first identify the solid and void intensity by fitting the pixel

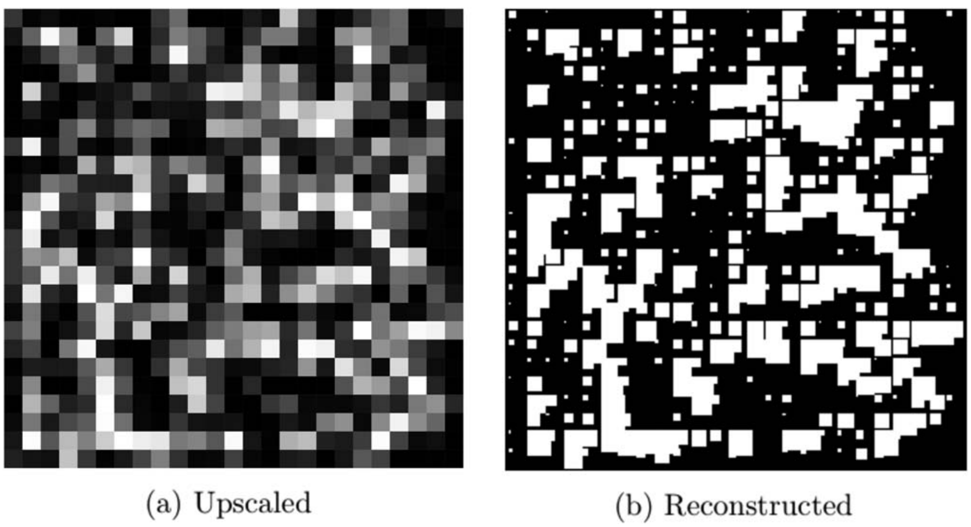


Figure 12. (a) Low-resolution upscaled image with added Gaussian noise, corresponding to the pore-scale distribution of Figure 10. The dimensions of the upscaled image are 10 × 10 pixels, i.e., 25 times smaller than the original high-resolution image. (b) Reconstructed pore-scale image using the downscaling algorithm. The dimensions of the reconstructed image are 250 × 250 pixels.

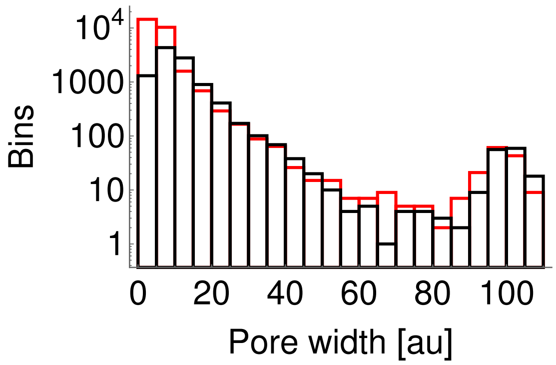


Figure 13. Comparison of histograms of the pore size distribution extracted from the original high-resolution (black lines) and reconstructed (red lines) images. The vertical axis has logarithmic scale.

intensity histogram with Gaussian functions as shown in Figure 18a. Then, by means of the intensity-porosity relation (2), we estimate the porosity of each image $\bar{\phi}_k$ as well as the total porosity of each column $\bar{\phi}$. For those images where a maximum corresponding to the void phase is not clearly evident, we use linear regression based on the extracted data to linearly correlate $v_v - v_s$ to v_s as shown in Figure 18b. The estimated value of the total porosity of stack $\mathcal{S}_1$ is $\bar{\phi}=0.178$, which perfectly matches the experimental value of $\phi=0.178$ for the column C6197A. For the image stack $\mathcal{S}_2$, the estimated porosity is $\bar{\phi}=0.278$. This value matches the experimentally measured value ($\phi=0.300$) reasonably well. We emphasize that the porosity estimate by means of the proposed new image processing algorithm (Step 1) does not have any fitting parameter, unlike segmentation-based approaches. This demonstrates the first step of the algorithm performs well on images of geologic media.

3.2.2. Step 2: Downscaling, Binary Image Reconstruction, and Segmentation

Once the porosity distribution is extracted from the grey scale image, a binary image, that preserves the connectivity of the original sample, is needed to run pore-scale flow and transport models. In this second step, we extract the pore size distribution of each image by applying both the physical and morphological models. In both cases the free parameter δ is determined such that the total porosity of the segmented image matches the porosity determined at Step 1 with a tolerance of E-5. Figure 19 shows the segmentation result after the downscaling step for both the physical and the morphological model, as well as a threshold-based segmentation. It is apparent that pore-space distribution is better captured by the morphological model. The dimension of the downscaled (high-resolution) images is $N_l \times N_{m,} = 20{,}480 \times 20{,}480$. To improve the execution time of the downscaling procedure, we parallelized the discretization algorithm using the OpenMP library (S. Korneev, https://github.com/svyatoslavkorneev/downscaling). We executed the parallel code on Amazon Web Services (AWS) compute-optimized EC2 36 core instance (Amazon Web Services data are available at https://aws.amazon.com/). The execution time for each image is approximately 2 min. For the binary image segmentation, we used the open-source library *scikit-image* (Image processing in Python data are available at http://scikit-image.org/).

3.2.3. Steps 3–4: From Processed Images to Flow and Transport Predictions—A Capillary-Bundle Model

The pore size distribution (extracted from either threshold-based segmentation or from the downscaling approach) can be directly used (i) to parametrize flow and transport models and (ii) to determine permeabilities. The permeabilities predicted from a given pore size distribution can then be compared with the experimental ones. For such estimates, we use a capillary-bundle model since pore-scale simulations of the reconstructed highly resolution images are computationally too expensive. We emphasize that, while any other approach can be employed to estimate flow and transport dynamics from a reconstructed binary image (including, but not limited to, direct numerical simulations), the aim of this study is to verify the performance of the image processing method with downscaling against that of a classic threshold-based segmentation approach: this is achieved by using the same capillary-bundle model on binary images obtained from either threshold-based or downscaling-based segmentation to determine conductivity estimates.

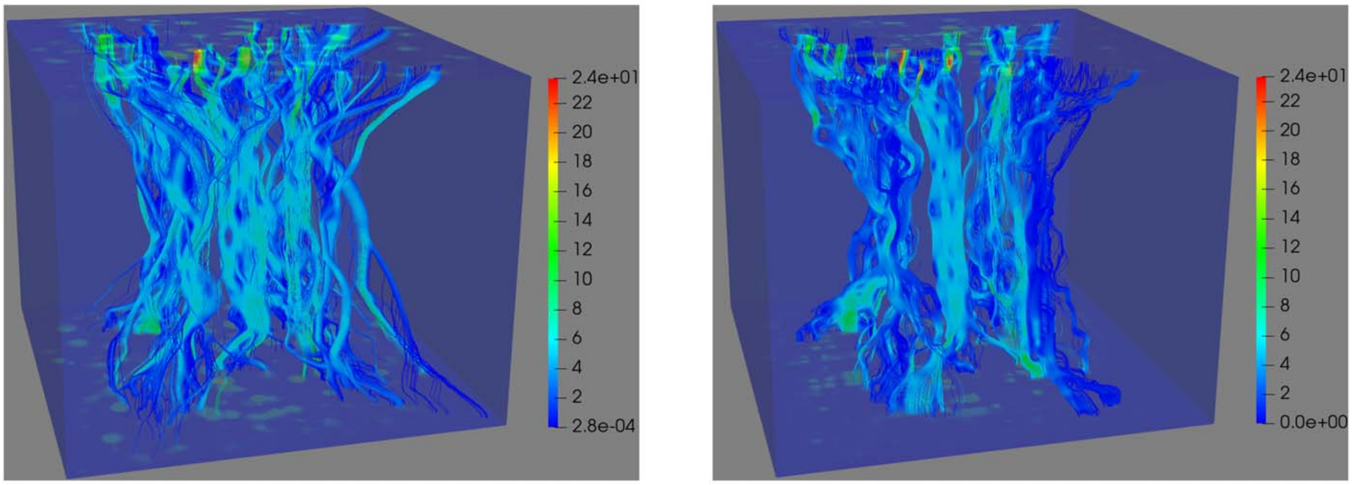
Here we idealize both experimental columns as bundles of capillary tubes of different radii, randomly blocked by transverse low permeable inclusions and embedded in a low-conductivity matrix (see cartoon in Figure 20). We also assume that the blockages/inclusions conductivity inside the capillary tubes, $k_{\text{inclusion}}$, is the same as that of the matrix which the capillaries are embedded in, k_{matrix}, i.e., $k_{\text{inclusion}} = k_{\text{matrix}}$. This idealization is supported by pore-scale simulations run by Scheibe et al. (2015b).

Once the image is segmented (using either threshold-based or downscaling-based segmentation) and the pore size distribution calculated, we rank the area of the extracted pores in image j in descending order, for any $j = 1, \ldots, N_k$, i.e.,

Table 2
Conductivity and Tortuosity of the Original and Reconstructed Domains

	K_z	τ_z
Original	1	1.2
Reconstructed	0.8	1.2

Note. The vertical conductivity of the samples is normalized by the conductivity of the original pore-scale domain. Tortuosity is calculated directly from estimating the normalized length streamlines in Figure 14.



(a) Original (b) Reconstructed

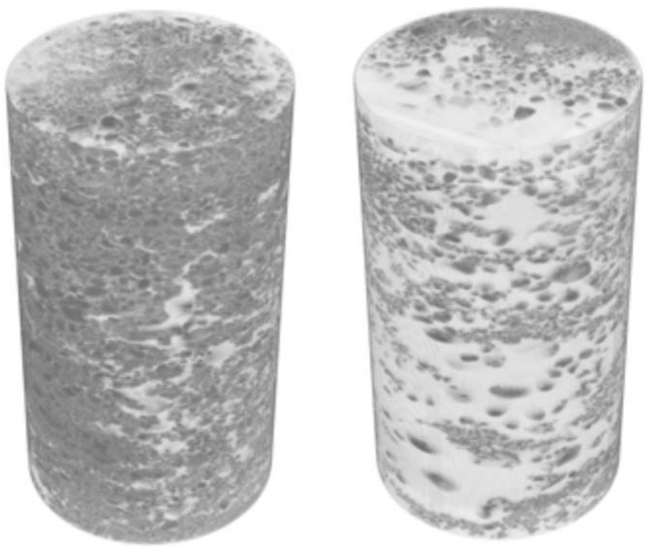
Figure 14. Streamlines topology and velocity magnitude (color legend) in the (a) original domain and (b) reconstructed domains. The transparent cube corresponds to the boundaries of the two media.

$$\begin{gathered}
\{s_1^1 > s_2^1 > s_3^1 > \cdots > s_c^1 > \cdots > s_p^1 > \bar{s}^1_{\text{matrix}}\} \text{ for image } 1 \\
[\ldots] \\
\{s_1^j > s_2^j > s_3^j > \cdots > s_c^j > \cdots > s_p^j > \bar{s}^j_{\text{matrix}}\} \text{ for image } j \quad , \\
[\ldots] \\
\{s_1^{N_k} > s_2^{N_k} > s_3^{N_k} > \cdots > s_c^{p} > \cdots > s_{N_c}^{N_k} > \bar{s}^{N_k}_{\text{matrix}}\} \text{ for image } N_k
\end{gathered} \tag{9}$$

where s_c^j is the c^{th} pore area (measured in pixels) of the j^{th} image, and s_p^j represents the pore size threshold in image j which separates pores constituting the matrix from pores forming the capillaries, i.e., the c^{th} pore is assigned to the *matrix* if $c > p$, and to the *capillary bundle*, otherwise. Also, $\bar{s}^j_{\text{matrix}}$ is the average area for pores with index c such that $p < c \le N^j$, where N^j is the total number of pores in image j, i.e., $\bar{s}^j_{\text{matrix}} := [\sum_{c=p+1}^{N^j} s_c^j]/(N^j - p - 1)$. The average pore radius for the matrix, $\bar{r}_{\text{matrix}}$, can be determined as $\bar{r}_{\text{matrix}} = \Delta\sqrt{\langle \bar{s}^j_{\text{matrix}} \rangle / \pi}$ where $\langle \bar{s}^j_{\text{matrix}} \rangle$ is the matrix pore area averaged across all images N_k and $\Delta = \Delta x = \Delta y$ is the dimensional pixel's size. Assuming that the matrix is itself a bundle of parallel capillary tubes with characteristic radius $\bar{r}_{\text{matrix}}$, the volumetric flux through the matrix Q_{matrix} can be estimated as

$$Q_{\text{matrix}} = N_{\text{matrix}} \frac{\pi \bar{r}^4_{\text{matrix}}}{8\mu} \Delta p, \tag{10}$$

where N_{matrix} is the number of capillaries inside the matrix. Since the total matrix cross section is



(a) C6197A (b) C6203A

Figure 15. Volume render from XCT images of columns (a) C6197A and (b) C6203A.

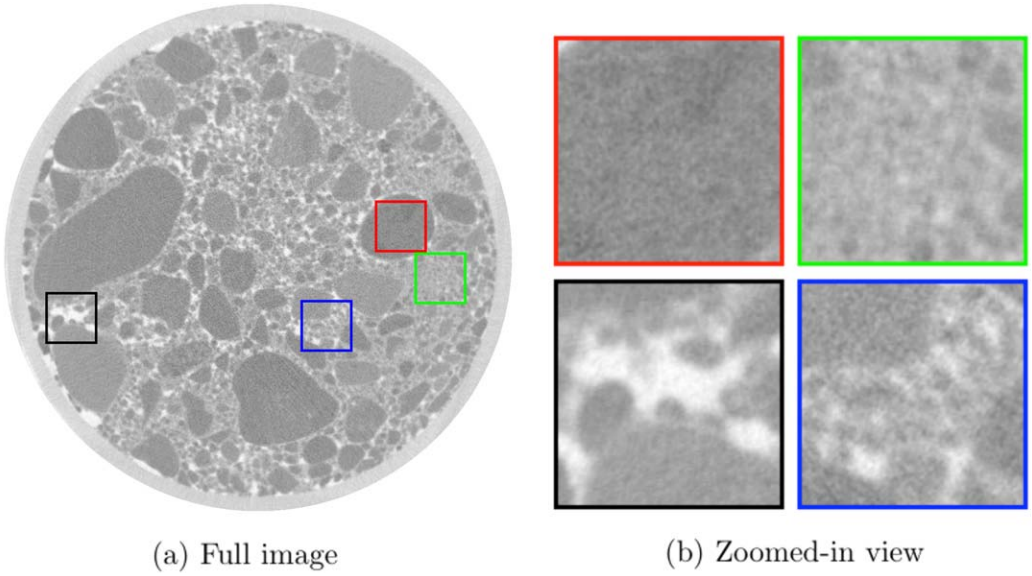


Figure 16. (a) A representative XCT image of sediment column C6197A. (b) Zoomed-in views of the solid phase (red square), porous-solid phase (green square), void phase (black square), and a mixture of the three phases (blue square).

$$A_{\text{matrix}}=\frac{\pi N_{\text{matrix}}\bar{r}^2_{\text{matrix}}}{\bar{\phi}}, \tag{11}$$

the matrix conductivity $[L/T]$ is

$$K_{\text{matrix}}=\rho g\bar{\phi}\frac{\bar{r}^2_{\text{matrix}}}{8\mu}. \tag{12}$$

Assuming the porosity of the inclusions (inside the capillaries) is equal to that of the matrix, N_{matrix} can be estimated as follows:

$$N_{\text{matrix}}=\frac{R^2-\sum_{c=1}^{p} r_c^2}{\bar{r}^2_{\text{matrix}}}\bar{\phi}, \tag{13}$$

where R is the physical radius of the column and $\bar{\phi}$ is the mean core porosity.

Each capillary c, with $c=1,\ldots,p$, is constructed by connecting pores with the same subscript between adjacent images and is assigned an effective radius r_c determined as the radius of the average pore area $\langle s_c^j\rangle$ across all N_k images, i.e., $r_c :=\Delta\sqrt{\langle s_j\rangle/\pi}$. Assuming that the p capillary tubes with radii $\{r_1, r_2, \ldots, r_p\}$ contain low conductivity inclusions with characteristic pore radius $\bar{r}_{\text{matrix}}$, then each clogged capillary c is a layered system and its volumetric flux can be estimated as follows:

$$Q_c=\frac{1}{\mu}\frac{L}{\frac{8(L-l)}{\pi r_c^4}+\frac{8l}{(r_c^2\bar{\phi}/\bar{r}^2_{\text{matrix}})\pi\bar{r}^4_{\text{matrix}}}}\Delta p, \tag{14}$$

where l is the length of the blockage and $r_c^2\bar{\phi}/\bar{r}^2_{\text{matrix}}$ the number of the capillary tubes inside the inclusion. Similarly, the conductivity of each capillary tube is

$$K_c=\rho g\frac{Q_c}{\pi r_c^2\Delta p}. \tag{15}$$

The equivalent conductivity for the bundle of c capillary tubes, K_{bundle}, can be defined as

$$K_{\text{bundle}}=\frac{\sum_{c=1}^{p} r_c^2 K_c}{\sum_{c=1}^{p} r_c^2}. \tag{16}$$

The total conductivity of the system is

Table 3
Experimental Measurements

	C6197A	C6203A
ϕ	0.178	0.300
K (cm/min)	1.570	1.956
Q (cm^3/min)	3.24	3.24
L (cm)	25	25
R (cm)	4.45	4.45
N_k	1,000	1,000
$N_i \times N_j$ (pixels^2)	1,024 × 1,024	1,024 × 1,024
$\Delta x=\Delta y$ (cm)	9.3×10^{-3}	9.3×10^{-3}
Δz (cm)	2.5×10^{-2}	2.5×10^{-2}
D (cm^2 min^{-1})	1.248×10^{-3}	1.248×10^{-3}
μ_w (Pa s)	1.002×10^{-3}	1.002×10^{-3}
ρ_w (kg m^{-3})	998.2	998.2
g (m s^{-2})	9.8	9.8

10.1002/2018WR022886

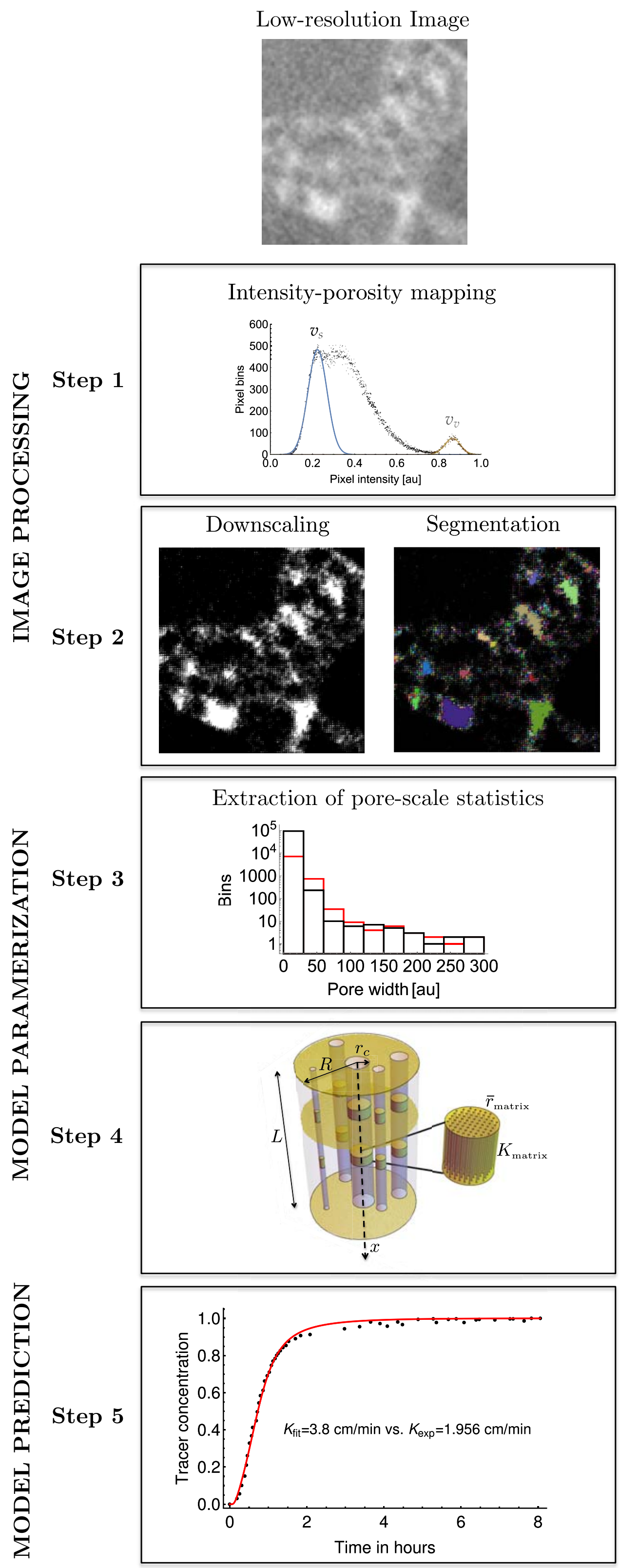


Figure 17. Image processing algorithm and its coupling to a physical model for flow and transport. The segmentation step in the image processing allows one to extract relevant pore-scale statistics, which are then used to parametrize a physical model of flow and transport (in this case, a capillary-bundle model).

$$K=\frac{A_{\mathrm{matrix}}K_{\mathrm{matrix}}+\pi\sum_{c=1}^{p}r_c^2K_c}{\pi R^2}, \tag{17}$$

where A_{matrix} is defined by (11), since $\pi R^2=A_{\mathrm{matrix}}+\pi\sum_{c=1}^{p}r_c^2$ (from (13)). Once K is determined, the concentration $c(t)$ of a passive solute measured at the end of the column $x=L$, is given by

$$c(t)=\frac{1}{Q}\left[\sum_{c=1}^{p}Q_c c_c(L,t)+Q_{\mathrm{matrix}}c_{\mathrm{matrix}}(L,t)\right], \tag{18}$$

where Q is the total volumetric flux through the column, and c_c and c_{matrix} are a solution of the 1-D transport equation:

$$\frac{\partial c_i}{\partial t}=D_i^\star\frac{\partial^2 c_i}{\partial x^2}-v_i\frac{\partial c_i}{\partial x},\quad i=\{c,\mathrm{matrix}\} \tag{19}$$

subject to

$$c_i(0,\,t)=c_{\mathrm{in}}\quad\text{and}\quad\frac{\partial c_i}{\partial x}(L,t)=0. \tag{20}$$

In (19),

$$v_i=\frac{K_i}{\rho g}\Delta p,\quad i=\{c,\mathrm{matrix}\}, \tag{21}$$

where K_{matrix} and K_c are defined by (12) and (15), respectively,

$$\Delta p=\frac{\rho g Q}{\pi R^2 K}, \tag{22}$$

and $D_i^\star$ is the Taylor dispersion coefficient defined as (Aris, 1956, equation (26))

$$\frac{D_i^\star}{D}=1+\frac{\mathrm{Pe}_i^2}{48}, \tag{23}$$

with

$$\mathrm{Pe}_i=\frac{r_i v_i}{D}, \tag{24}$$

and D the molecular diffusion coefficient. Equation (19) has a known analytical solution:

$$c_i(x,\,t)=\frac{1}{2}\mathrm{erfc}\left(\frac{x-v_it}{2\sqrt{D_i^\star t}}\right)+\frac{1}{2}\exp\left(\frac{v_ix}{D_i^\star}\right)\mathrm{erfc}\left(\frac{x+v_it}{2\sqrt{D_i^\star t}}\right), \tag{25}$$

with $i=\{c,\mathrm{matrix}\}$. In the following, we compare the model predictions (17) and (18) with experimentally measured conductivity values and breakthrough curves.

3.2.4. Step 5: Comparison Between Model Predictions and Data Using Threshold-Based and Downscaling-Based Segmentation

Since on the experiment time scale t_{end} (<8h), $Q_{\mathrm{matrix}}c_{\mathrm{matrix}}(L,t)\ll\sum_{c=1}^{p}Q_cc_c(L,t)$, then (18) can be simplified as follows:

$$c(t)\approx\frac{1}{Q}\sum_{c=1}^{p}Q_cc_c(L,t). \tag{26}$$

This hypothesis will be tested a posteriori. The breakthrough curve (26) contains two unknown parameters: the number of capillary tubes p and the length l of the inclusions. We determine p and l by minimizing the L_2 norm:

10.1002/2018WR022886

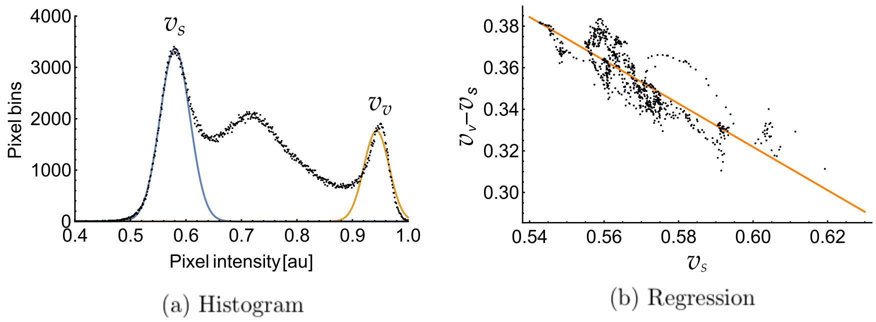


Figure 18. (a) Pixel intensity histogram of a XCT image (black dots) and the Gaussian fits corresponding to the solid and void phases (blue and orange solid lines, respectively). (b) Linear regression (orange line) between $v_v - v_s$ and v_s (black dots).

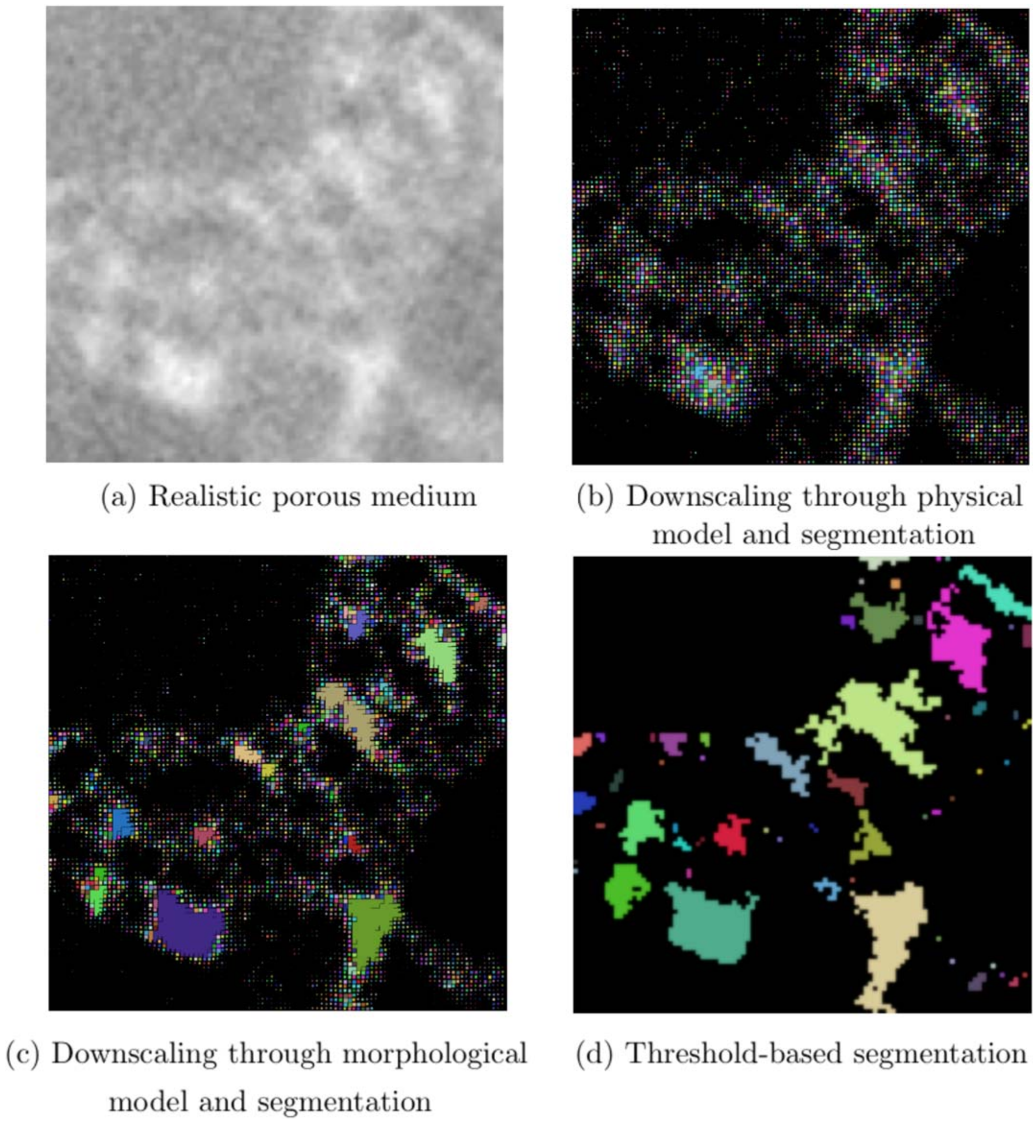


Figure 19. Example of segmentation with downscaling of a (a) real image, where each color labels an interconnected bounded region (i.e., a pore). (b) Segmentation with downscaling through the physical model. (c) Segmentation with downscaling through the morphological model. (d) Threshold-based segmentation.

10.1002/2018WR022886

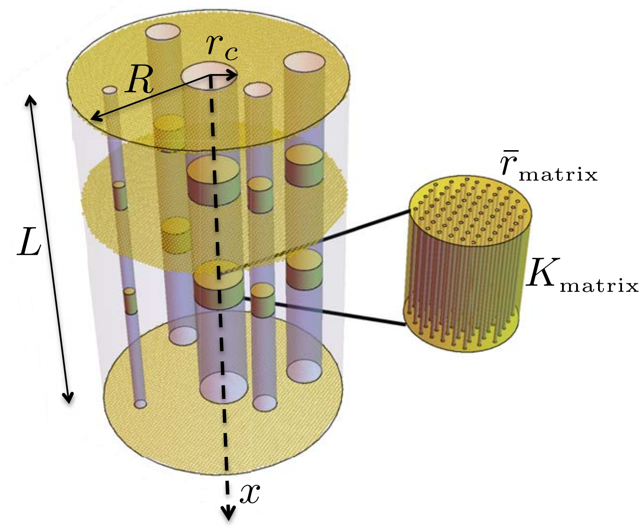
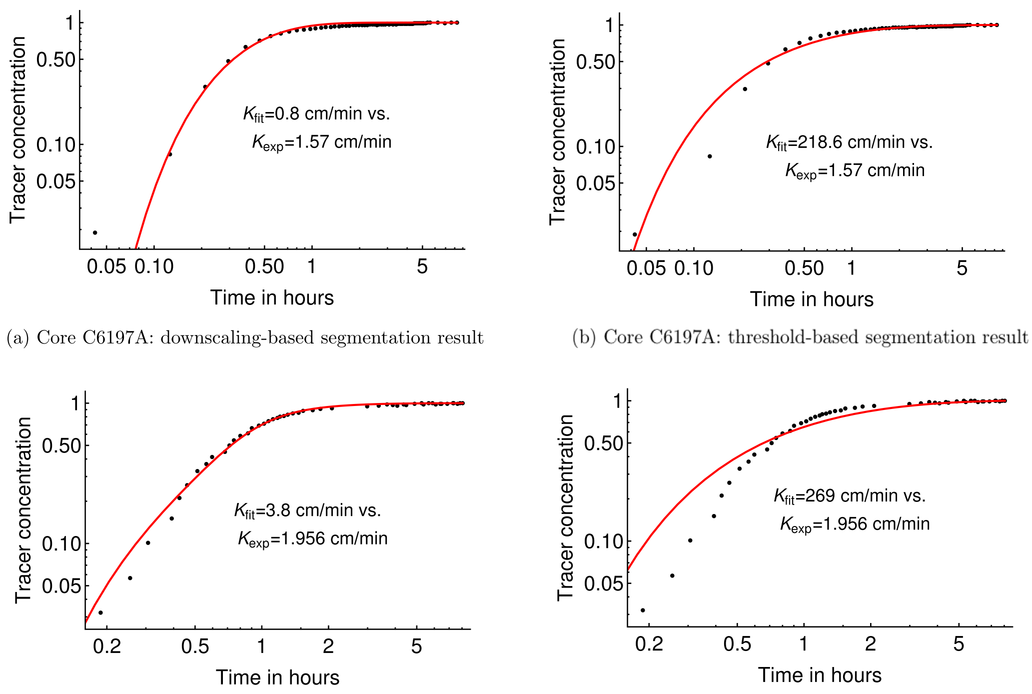


Figure 20. Idealization of the experimental core as a bundle of capillaries embedded in a porous matrix of conductivity K_{matrix}, whose pores have a uniform characteristic size $\bar{r}_{\text{matrix}}$. Each capillary c has radius r_c and contains a low conductivity inclusion. The latter has the same properties of the surrounding matrix.

$$L_2 = \left\{ \sum_{i=1}^{N_b} \left[\frac{c(t_i)[p, l]}{c(t_{\text{end}})} - b_i \right]^2 \right\}^{1/2}, \tag{27}$$

where $\boldsymbol{b_i}$ is the experimental concentration measured at time t_i, normalized by b_{N_b}, and $t_{N_b} = t_{\text{end}}$. We emphasize that, while we use p and l (instead of a more classical dispersion coefficient D_{eff}) as fitting parameters to minimize the error between the theoretical breakthrough curve (26) and the experimental one, the dependence of the BCT on D_{eff} is implicit: the effective dispersion coefficient of the core depends on the number p of capillaries in the bundle (as evident from equation (26)), as well as on the length l of the low conductivity inclusions within them. In Table 3, we list the conductivity measurements for both cores as well as other experimental quantities. Once p and l are defined, we determine K from (17). For each core, we run the minimization problem on the binary images generated from a threshold-based or downscaling-based (with morphological model) segmentation approach. It is worth noticing that in the threshold-based method each image is segmented such that its porosity matches that of the original grey scale image. Figure 21 shows the fitted breakthrough curves for the two cores C6197A and C6203A. While both image processing methods provide a reasonably good fit of the breakthrough curves, the conductivity predicted from threshold-based segmentation overestimates the experimenatlly measured conductivity by 2 order of magnitudes for both samples. The downscaling-based segmentation provides an



(a) Core C6197A: downscaling-based segmentation result

(b) Core C6197A: threshold-based segmentation result

(c) Core C6203A: downscaling-based segmentation result

(d) Core C6203A: threshold-based segmentation result

Figure 21. Comparison between experimental and fitted breakthrough curves from different image processing methods (downscaling-based on the left, and threshold-based on the right). Predicted conductivity values are also compared with experimental ones. Despite the breakthrough curves can be fitted with either method, a threshold-based segmentation approach provides conductivity predictions that are 2 order of magnitude larger than the experimentally measured values. (a) Core C6197A: downscaling-based segmentation result. (b) Core C6197A: threshold-based segmentation result. (c) Core C6203A: downscaling-based segmentation result. (d) Core C6203A: threshold-based segmentation result.

Table 4
Parameters of the Capillary-Bundle Model for Both Cores C6197A and C6203A

	C6197A		C6203A	
	Downscaling based	Threshold based	Downscaling based	Threshold based
L_2 norm	0.26	0.24	0.13	0.44
l	1	0.5	0.75	1.5
p	1	1	6	1
K (cm/min)	**0.8**	**218.8**	**3.8**	**269.0**
ϕ	**0.178**	**0.178**	**0.277**	**0.277**
K_{matrix} (cm/min)	0.63	63.2	1.0	108.1
K_{bundle} (cm/min)	15.7	3,149.1	34.6	1,800.3
$\bar{r}_{matrix}$ (cm)	2.2×10^{-3}	2.4×10^{-2}	2.3×10^{-3}	2.5×10^{-2}
Q_{matrix} (cm^3/min)	2.48	0.89	0.8	1.2
v_{matrix} (cm/min)	4.0×10^{-2}	1.51×10^{-2}	1.4×10^{-2}	2.1×10^{-2}
Pe_{matrix}	0.14	0.51	0.05	0.75
$\tau_{matrix}=L/v_{matrix}$ (min)	618.4	1,659.2	1.783.3	1,193.7
Q_{bundle} (cm^3/min)	0.75	2.35	2.4	2.1
r_1 (cm)	4.9×10^{-1}	1.0	9.4×10^{-1}	1.4
v_1 (cm/min)	1	0.75	0.47	0.35
Pe_1	760.6	1,156.0	675.0	738.0
$\tau_1=L/v_1$ (min)	24.7	33.3	53.5	71.7

Note. All the parameters are determined after fitting p and l by minimizing the L_2 norm (27). The predicted conductivity and porosity using either threshold-based or downscaling-based segmentation approaches are in boldface.

excellent conductivity estimate (see Table 4). This result can be explained as follows: while threshold-based segmentation can capture the characteristic size of the largest pores (i.e., the capillary bundle), it fails to properly describe the characteristic size of the unresolved pore spaces. Since each capillary acts as a in-series layered system, its effective conductivity is primarily controlled by the smallest pores, see equation (14). As a result, the conductivity predicted after binary segmentation significantly differs from the experimental one despite the breakthrough curves can be fitted reasonably well. A similar result was obtained by Scheibe et al. (2015b), where the estimated conductivity of the reconstructed core after a binary segmentation was 10 times larger than the experimental value. In that case, also the estimated porosity differed by 1 order of magnitude from the measured one (see their Table 1). Table 4 has a list of characteristic quantities determined from both downscaling-based and threshold-based segmentation. While the characteristic quantities of the capillary bundle (e.g., K_{bundle}) and the largest capillary (e.g., radius r_1, Darcy flux v_1, Peclét number Pe_1, advection time through the column τ_1) obtained from either image segmentation method have the same order of magnitude, threshold-based segmentation leads to an estimate of the characteristic radius of the matrix r_{matrix} that is 1 order of magnitude larger than that from the downscaling-based method. This results in a significant overestimation of the conductivity K of the cores by 2 order of magnitudes. Finally, we determine the advection time scale for the matrix to test the approximation (26). Since for all cases, the advection time inside the matrix is more than 8 h (as shown in Table 4), it is reasonable to neglect the matrix contribution to the breakthrough curves.

4. Conclusions

Notwithstanding the advancements in direct numerical simulations of pore-scale flow and transport in natural sediments, the accurate reconstruction of pore-scale topology from X-ray tomographic images remains a major bottleneck in the development of predictive computational models. The limited resolution of X-ray tomography is often insufficient to resolve and characterize the geometry of the pore space within the fine-grained region of a sample. As a result, an overwhelming number of studies have used single and multilevel threshold-based segmentation approaches to identify different phases in the sample (generally, a pore, a solid and a porous-solid phase) and reconstruct the pore-scale topology necessary to run direct numerical simulations. Such studies have shown that the dynamical response of the reconstructed porous structure strongly depends on the choice of such (arbitrary) cutoff thresholds. Here we have proposed a threshold-free downscaling-based segmentation approach to handle partial volume effects in moderately unresolved images of spatially heterogeneous, but chemically homogeneous, granular media. The method is based on (i) a map between pixel grey scale intensity and pixel porosity and (ii) a subpixel morphological model used to generate a high-resolution binary image from poorly resolved porosity maps while preserving important subresolution information and intraslice connectivity of large pores. The primary advantage of the proposed algorithm is its ability to better capture relevant unresolved lengths scales without relying on the definition of thresholds. The downscaling-based reconstruction algorithm is first tested on synthetic images and its performance compared with threshold-based segmentation approaches: specifically, we show that the downscaling-based method is able to represent the full pore size distribution, including below-resolution length scales, better than threshold-based segmentation. Next, we test the algorithm on a synthetic 3-D porous medium and show that the reconstructed image can be successfully used to calculate effective transport properties (e.g., conductivity and tortuosity) through Direct Numerical Simulations. We also apply the algorithm on XCT images of two sediment columns collected from boreholes in the 300 Area of the U.S. Department of Energy Hanford Site in South-Central Washington State. We generate binary images of the pore structure from the original unresolved XCT images by means of both threshold-based and downscaling-based segmentation. Once the binary images are reconstructed, we use them to parametrize an idealized capillary-bundle model to describe the topology of both columns. While the predicted conductivity from the threshold-based segmentation overestimates the experimental value by 2 orders of magnitude, the downscaling-based segmentation provides an excellent prediction due to its ability to better characterize the length scale distribution of small pores. Finally, while the algorithm is designed to process moderately unresolved images of chemically homogeneous granular media with large spatial heterogeneity, the extension of the algorithm to chemically heterogeneous samples is subject of current investigations.

Acknowledgments
IB and SK have been supported by the U.S. Department of Energy Early Career Award in Basic Energy Sciences DE-SC0014227 and by the National Science Foundation under award number EAR-1742569. The scientific contributions from each grant are as follows: the image processing tool development and the image processing applied to Hanford cores was supported by the DOE Early Career Award, while the capillary-bundle model development (section 3.2: Step 3) was supported by the National Science Foundation. We are also grateful to Amazon AWS Programs for Research and Education that provided allocation in their cloud infrastructure. Data collection from the Hanford site was supported by the U.S. Department of Energy, Office of Biological and Environmental Research (BER) through the Subsurface Biogeochemical Research (SBR) Science Focus Area (SFA) project at Pacific Northwest National Laboratory (PNNL). The data are available at https://www.digitalrocksportal.org/projects/137. Authors contributions are as follows: Svyatoslaav Korneev developed the numerical and analytical models, processed the data, analyzed the results and wrote the manuscript. Xiaofan Yang assisted with the Hanford site data processing and interpretation, and provided feedbacks on the text. John Zachara and Timothy Scheibe provided the experimental dataset from the Hanford Site, as well as comments and feedbacks on the text. Ilenia Battiato conceptualized the study, integrated the results obtained from experiments, numerical simulations and the analytical models, and wrote the manuscript.

References

Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., et al. (2013a). Digital rock physics benchmarks—Part I: Imaging and segmentation. *Computers & Geosciences*, *50*, 25–32. https://doi.org/10.1016/j.cageo.2012.09.005

Andrä, H., Combaret, N., Dvorkin, J., Glatt, E., Han, J., Kabel, M., et al. (2013b). Digital rock physics benchmarks—Part II: Computing effective properties. *Computers & Geosciences*, *50*, 33–43. https://doi.org/10.1016/j.cageo.2012.09.008

Aoud, M. (2014). Image restoration based on morphological operations. *International Journal of Computer Science, Engineering and Information Technology*, *4*(3), 9–21. https://doi.org/10.5121/ijcseit.2014.4302

Aris, R. (1956). On the dispersion of a solute in a fluid flowing through a tube. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *235*(1200), 67–77.

Arunachalam, H., Onori, S., & Battiato, I. (2015). On veracity of macroscopic lithium-ion battery models. *Journal of the Electrochemical Society*, *162*(10), A1940–A1951. https://doi.org/10.1149/2.0771509jes

Battiato, I. (2014). Effective medium theory for drag-reducing micro-patterned surfaces in turbulent flows. *The European Physical Journal E: Soft Matter and Biological Physics*, *37*, 19.

Battiato, I. (2016). Multiscale models of flow and transport. In Cushman, J. H., & Tartakovsky, D. M. (Eds.), *Handbook of groundwater engineering* (Chapter 29) Boca Raton, FL: CRC Press.

Battiato, I., Tartakovsky, D., Tartakovsky, A., & Scheibe, T. (2009). On breakdown of macroscopic models of mixing-controlled heterogeneous reactions in porous media. *Advances in Water Resources*, *32*(11), 1664–1673. https://doi.org/10.1016/j.advwatres.2009.08.008

Battiato, I., & Tartakovsky, D. M. (2011). Applicability regimes for macroscopic models of reactive transport in porous media. *Journal of Contaminant Hydrology*, *120–121*(C), 18–26. https://doi.org/10.1016/j.jconhyd.2010.05.005

Battiato, I., Tartakovsky, D. M., Tartakovsky, A. M., & Scheibe, T. D. (2011). Hybrid models of reactive transport in porous and fractured media. *Advances in Water Resources*, *34*(9), 1140–1150. https://doi.org/10.1016/j.advwatres.2011.01.012

Bernabe, Y., Li, M., & Maineult, A. (2010). Permeability and pore connectivity: A new model based on network simulations. *Journal of Geophysical Research*, *115*, B10203. https://doi.org/10.1029/2010JB007444

Boas, F. E., & Fleischmann, D. (2012). CT artifacts: Causes and reduction techniques. *Imaging in Medicine*, *4*(2), 229–240. https://doi.org/10.2217/iim.12.13

Boso, F., & Battiato, I. (2013). Homogenizability conditions for multicomponent reactive transport. *Advances in Water Resources*, *62*, 254–265. https://doi.org/10.1016/j.advwatres.2013.07.014

Bultreys, T., De Boever, W., & Cnudde, V. (2016). Imaging and image-based fluid transport modeling at the pore scale in geological materials: A practical introduction to the current state-of-the-art. *Earth-Science Reviews*, *155*, 93–128. https://doi.org/10.1016/j.earscirev.2016.02.001

Bultreys, T., Van Hoorebeke, L., & Cnudde, V. (2014). Multi-scale, micro-computed tomography-based pore network models to simulate drainage in heterogeneous rocks. *Advances in Water Resources*, *78*, 36–49.

Davies, R., & Kasper, M. (2012). Adaptive optics for astronomy. *Annual Review of Astronomy and Astrophysics*, *50*(1), 305–351. https://doi.org/10.1146/annurev-astro-081811-125447

Fister, T. T., Seidler, G. T., Rehr, J. J., Kas, J. J., Elam, W. T., Cross, J. O., et al. (2007). Deconvolving instrumental and intrinsic broadening in core-shell x-ray spectroscopies. *Physical Review B*, *75*(17), 174106. https://doi.org/10.1103/PhysRevB.75.174106

Gerke, K. M., Karsanina, M. V., & Mallants, D. (2015). Universal stochastic multiscale image fusion: An example application for shale rock. *Scientific Reports*, *5*, 15880. https://doi.org/10.1038/srep15880

Hwang, S., & Felix, W. W. (2002). Subvoxel processing: A method for reducing partial volume blurring with application to in vivo MR images of trabecular bone. *Magnetic Resonance in Medicine*, *47*(5), 948–957.

Iassonov, P., Gebrenegus, T., & Tuller, M. (2009). Segmentation of X-ray computed tomography images of porous materials: A crucial step for characterization and quantitative analysis of pore structures. *Water Resources Research*, *45*, W09415. https://doi.org/10.1029/2009WR008087

Kaestner, A., Lehmann, E., & Stampanoni, M. (2008). Imaging and image processing in porous media research. *Advances in Water Resources*, *31*(9), 1174–1187. https://doi.org/10.1016/j.advwatres.2008.01.022

Korneev, S., & Battiato, I. (2016). Sequential homogenization of reactive transport in polydisperse porous media. *Multiscale Modeling & Simulation*, *14*(4), 1301–1318. https://doi.org/10.1137/16M1074278

Lei, T. (2012). *Statistics of medical imaging*. Retrieved from https://doi.org/10.2214/AJR.12.8952

Lindquist, W. B., Lee, S.-M., Coker, D. A., Jones, K. W., & Spanne, P. (1996). Medial axis analysis of void structure in three-dimensional tomographic images of porous media. *Journal of Geophysical Research*, *101*(B4), 8297–8310. https://doi.org/10.1029/95JB03039

Ling, B., Oostrom, M., Battiato, I., & Tartakovsky, A. M. (2017). Modeling variability in porescale multiphase flow experiments. *Advances in Water Resources*, *105*, 29–38.

Ling, B., Tartakovsky, A. M., & Battiato, I. (2016). Dispersion controlled by permeable surfaces: Surface properties and scaling. *Journal of Fluid Mechanics*, *801*, 13–42.

Mehmani, Y., & Tchelepi, H. A. (2017). Minimum requirements for predictive pore-network modeling of solute transport in micromodels. *Advances in Water Resources*, *108*, 83–98.

Mohebi, A., & Fieguth, P. (2008). *Statistical fusion and sampling of scientific images*. Paper presented at the Proceedings of the International Conference on Image Processing, ICIP (pp. 1312–1315). https://doi.org/10.1109/ICIP.2008.4712004

Mohebi, A., Fieguth, P., & Ioannidis, M. A. (2009). Statistical fusion of two-scale images of porous media. *Advances in Water Resources*, *32*(11), 1567–1579. https://doi.org/10.1016/j.advwatres.2009.08.005

Prodanović, M., Mehmani, A., & Sheppard, A. (2014). Imaged-based multiscale network modelling of microporosity in carbonates. *Geological Society, Special Publications*, *406*(2012), SP406–9. https://doi.org/10.1144/SP406.9

Samet, H., & Tamminen, M. K. (1988). Efficient component labeling of images of arbitrary dimension represented by linear bintrees. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *10*(4), 579–586. https://doi.org/10.1109/34.3918

Scheibe, T. D., Murphy, E. M., Chen, X., Rice, A. K., Carroll, K. C., Palmer, B. J., et al. (2015a). An analysis platform for multiscale hydrogeologic modeling with emphasis on hybrid multiscale methods. *Ground Water*, *53*(1), 38–56. https://doi.org/10.1111/gwat.12179

Scheibe, T. D., Perkins, W. A., Richmond, M. C., Mckinley, M. I., Romero-Gomez, P. D. J., Oostrom, M., et al. (2015b). Pore-scale and multiscale numerical simulation of flow and transport in a laboratory-scale column, *Water Resources Research*, *51*, 1023–1035. https://doi.org/10.1002/2014WR015959

Sok, R. M., Varrdslot, T., Ghous, A., Latham, S., Sheppard, A. P., & Knackstedt, M. A. (2010). Pore scale characterization of carbonates at multiple scales: Integration of micro-CT, BSEM, and FIBSEM. *Petrophysics*, *51*(6), 379–387.

Soulaine, C., Gjetvaj, F., Garing, C., Roman, S., Russian, A., Gouze, P., et al. (2016). The impact of sub-resolution porosity of X-ray microtomography images on the permeability. *Transport in Porous Media*, *113*(1), 1–17. https://doi.org/10.1007/s11242-016-0690-2

Sun, T., Mehmani, Y., & Balhoff, M. T. (2012). Hybrid multiscale modeling through direct substitution of pore-scale models into near-well reservoir simulators. *Energy and Fuels*, *26*(9), 5828–5836.

Wildenschild, D., & Sheppard, A. P. (2013). X-ray imaging and analysis techniques for quantifying pore-scale structure and processes in subsurface porous medium systems. *Advances in Water Resources*, *51*, 217–246. https://doi.org/10.1016/j.advwatres.2012.07.018

Yang, X., Mehmani, Y., Perkins, W. A., Pasquali, A., Schönherr, M., Kim, K., et al. (2016). Intercomparison of 3D pore-scale flow and solute transport simulation methods. *Advances in Water Resources*, *95*, 176–189.

Yoon, H., Kang, Q., & Valocchi, A. J. (2015). Lattice Boltzmann-based approaches for pore-scale reactive transport. *Reviews in Mineralogy & Geochemistry*, *80*, 393–431.

Yousefzadeh, M., & Battiato, I. (2017). Physics-based hybrid method for multiscale transport in porous media. *Journal of Computational Physics*, *344*, 320–338. https://doi.org/10.1016/j.jcp.2017.04.055

Zachara, J. M., Long, P. E., Bargar, J., Davis, J. A., Fox, P., Fredrickson, J. K., et al. (2013). Persistence of uranium groundwater plumes: Contrasting mechanisms at two DOE sites in the groundwater-river interaction zone. *Journal of Contaminant Hydrology*, *147*, 45–72. https://doi.org/10.1016/j.jconhyd.2013.02.001