Computational study of permeability in cardboard coating layers

Received: 29 September 2025 | Revised: 4 March 2026 | Accepted: 1 April 2026

DOI: 10.1002/aic.70388

AIChE JOURNAL

RESEARCH ARTICLE

Soft Matter: Synthesis, Processing and Products

Computational study of permeability in cardboard coating layers

Sandra Barman [1] | Torben Nilsson Pingel [1] | Niklas Lorén [1,2] | Tobias Gebäck [3] | Eva Olsson [2] | Aila Särkkä [3] | Magnus Röding [1,3,4]

[1]RISE Research Institutes of Sweden, Bioeconomy and Health, Agriculture and Food, Göteborg, Sweden

[2]Department of Physics, Chalmers University of Technology, Göteborg, Sweden

[3]Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, Göteborg, Sweden

[4]Sustainable Innovation & Transformational Excellence, Pharmaceutical Technology & Development, Operations, AstraZeneca Gothenburg, Mölndal, Sweden

Correspondence
Sandra Barman, RISE Research Institutes of Sweden, Frans Perssons väg 6, 41276 Göteborg, Sweden.
Email: sandra.barman@ri.se

Funding information
VINNOVA, Grant/Award Number: 2021-03665; Vetenskapsrådet, Grant/Award Numbers: 2022-06725, 2018-05973

Abstract

We develop a virtual material structure model based on a combination of tessellations and Gaussian random fields for a coating layer of paperboard used for packaging and designed to facilitate printing on the surface. To fit the model to tomographic image data acquired using combined focused ion beam and scanning electron microscopy (FIB-SEM), we introduce a novel method for estimating a covariance function, combined with approximate Bayesian computation (ABC). We demonstrate good agreement with the real material in terms of several microstructural descriptors. We then use the developed model in a computational study to establish structure–property relationships, specifically how permeability varies as a function of porosity, length scale, and other parameters of the virtual structure model. We conclude that the variation in permeability can be explained very well by porosity and descriptors that capture path lengths through the pore system, bottleneck effects, and the specific surface area.

KEYWORDS

3D spatial geometry model, cardboard coating, microstructural descriptors, permeability, virtual materials testing

1 | INTRODUCTION

To optimize the functionality of porous materials, it is crucial to characterize the relationship between microstructure and mass transport properties. Here, we are interested in porous coatings used on paperboards to support high quality printing on the surface. The expected behavior of the coating is to retain ink pigments but let the solvent through. Detailed understanding of the microstructure requires high resolution imaging and given the small length scales, X-ray computed tomography is not a feasible technique. However, combined focused ion beam and scanning electron microscope (FIB-SEM) tomography can address these challenges.[1] Using the FIB-SEM technique, a three-dimensional volume is imaged as a sequence of two-dimensional cross-section images. The FIB is used to gradually mill away thin slices of the sample, and the SEM is then used to image the planar cross-sections. The process is repeated until a sufficient field of view is obtained. The porous microstructure can then be segmented and characterized.

Exploring microstructural parameters and their importance for mass transport using experimental characterization of a large number of samples is expensive and time-consuming. Therefore, it is beneficial to develop a statistical model for the microstructure and explore the parameter space of the model instead. Exploring virtual microstructure models has been used as a means of understanding

microstructure-property relationships across many domains including batteries,[2] polymer coatings for drug release,[3] and fiber materials.[4] This type of modeling can fit a broad range of materials, tweaking parameters to fit the model to a specific real microstructure. The modeling does not specifically mimic the physical processes creating the material but instead takes an empirical approach mimicking its microstructural features.

In this work, we develop a statistical model for a porous coating used on paperboards, based on a model for the same material that was investigated in the previous study.[5] The model is based on a combination of sphere packings, Laguerre tessellations, and Gaussian random fields, and the parameters of the model are estimated by combining a novel method for covariance function estimation with approximate Bayesian computation (ABC). Using the model, we generate structures with different porosities, length scales, and microstructural properties, and study which impact these characteristics have on fluid permeability. Finally, we fit additive models[6,7] to explain the relationship between fluid permeability and the microstructure by using the microstructural descriptors of the structure.

The paper is organized as follows. First, we describe the experimental data and characterize it by using different microstructural descriptors. Then, we introduce the model and describe the model fitting approach and assess its performance. Finally, we explore structures with different parameters and permeability in a simulation study and establish how permeability can be explained by the microstructure.

2 | RESULTS AND DISCUSSION

2.1 | Experimental data

We study a sample of a coating applied to paperboards intended for food packaging. The coating consists of a powder of calcium carbonate particles (80%) and a latex-based binder (20%). The percentages are specified relative to the total solid volume and hence do not account for the porosity of the coating. Image data was acquired using combined focused ion beam and scanning electron microscope (FIB-SEM) tomography performed using a Tescan GAIA3 (Tescan, Brno, Czech Republic). Further, an Emitech K550X sputter coater (Quorum Technologies, Laughton, UK) was used to coat the sample with a thin palladium layer for increased electrical conductivity. The microscope is equipped with a backscattered electron detector and a gallium ion column. An ion beam energy of 30 keV and electron beam energy of 1.55 keV were used. The captured images have a pixel size in each SEM image of 10 nm, and the thickness of each slice removed by the FIB is 20 nm. After alignment and cropping in Fiji[8] and tomviz,[9] the size of the reconstructed 3D volume is 5.00 μm × 6.00 μm × 4.28 μm. More details are given in,[5] in which this sample is referred to as 'Sample A'.

Segmentation of the 3D volume, that is, separating the porous network from the solid, was performed using a machine learning method. Manual annotation of parts of the data, performed by an expert, was used to train a random forest-based classifier on features extracted from a set of Gaussian smoothing filters (linear scale-space features) applied to the data. Then, the algorithm was applied to the full dataset, followed by a smoothing to reduce classification noise. The validation accuracy was approximately 92%. More details are given in reference [5] and additionally in reference [10]. The segmentation was implemented in Matlab (Mathworks, Natick, MA, USA).

The experimental data is further illustrated in Section 2.2.

2.2 | Characterization

We use microstructural descriptors throughout the paper: to characterize the sample, to fit the model, to assess model fit, to explore the effect of changing model parameters on the geometry of model structures, and to study the relationship between the geometry and the permeability. Here we provide a description of the microstructural descriptors that are used, followed by 3D visualizations of a few of the descriptors computed for the experimental data.

2.2.1 | Microstructural descriptors

We first introduce some notation. Let $W = [0, L_x] \times [0, L_y] \times [0, L_z]$ be the observation window. The faces of the window W in a given direction are denoted inlet face F_{in} and outlet face F_{out}. The printing ink is transported through the pores in the x direction, which is therefore the most interesting direction both for microstructural characterization and flow simulations. We denote by $\mathcal{D} : W \to \{0, 1\}$ a binary field representing a structure with a pore phase and solid phase, such as the segmented data or a realization from the model presented in Section 2.4. $\mathcal{D}(\mathbf{x})$ then equals one if $\mathbf{x} = (x, y, z) \in W$ is a part of the pore space and zero otherwise. Second, we introduce seven microstructural descriptors that will be used in the remaining part of the paper. The most basic one is total porosity, which is simply the volume fraction of the pore phase, and effective porosity, which is the volume fraction of pores that are connected to both the inlet and outlet face in the x direction. Here, we use 26-connectivity to define which voxels are neighbors, and count a voxel as being connected to a subset of W if there is a path of connected pore voxels between the voxel and the set, in this case inlet and outlet.

The second descriptor is the specific surface area s of the interface between the pore phase and solid phase which correlates both with pore volume and smoothness of the pore walls. Details about methodology used for approximating a surface from a pixelized/voxelized set can be found in references [11,12].

The third and fourth descriptors, contact distance and 3D sphere size, are both defined with respect to a structuring element, in this case a sphere. The contact distance[13] is computed using sequential erosions with respect to the sphere with increasing size. Both erosion and opening below are morphological operators, see[13] for details. The contact distance of a point $\mathbf{x}$ belonging to a subset $\Gamma \subseteq W$, denoted $c_{sphere}(\mathbf{x})$, equals the radius of the largest sphere centered at $\mathbf{x}$ that can fit within Γ. The 3D sphere size, more commonly known as granulometry or 3D continuous pore size, is computed using sequential

openings with increasing size. The 3D sphere size at $\mathbf{x}$, denoted $\text{size}_{\text{sph}}(\mathbf{x})$, equals the size of the largest sphere that can fit within Γ and cover the point $\mathbf{x}$. Both descriptors are useful for defining sizes of complex objects. Since the contact distance can be efficiently computed using distance transforms, it is much faster to compute than the 3D sphere size. The 3D sphere size, on the other hand, is closer to the intuition of the size of an irregular object making it attractive in interpretation of size distributions and 3D size visualizations. Here, each descriptor is computed once in the pore phase, that is, with $\Gamma = \{\mathbf{x} : \mathcal{D}(\mathbf{x}) = 1\}$, and once in the solid phase, that is, with $\Gamma = \{\mathbf{x} : \mathcal{D}(\mathbf{x}) = 0\}$.

The fifth descriptor, through-intrusion porosimetry, emulates the experimental process of mercury intrusion porosimetry. Similar to the 3D sphere size, it is computed with sequential openings, but with the added step of computing the part of the opened structure that is connected to a chosen set of faces of the window W for each opening. We choose one of either the inlet F_{in} or the outlet F_{out} in the x direction as the starting face. For the case of F_{in}, the intrusion porosimetry size $\text{size}_{\text{IP},F_{\text{in}}}(\mathbf{x})$ in a point $\mathbf{x}$, relative to the subset Γ and face F_{in}, equals the size of the largest sphere that can travel from $F_{\text{in}} \cap \Gamma$ through Γ to a position where it covers the point $\mathbf{x}$. Just as with the 3D sphere size, the sphere needs to cover the point $\mathbf{x}$ but it does not need to be centered at $\mathbf{x}$. We then define the through-intrusion porosimetry as

$$\text{size}_{\text{IP,through}}(\mathbf{x}) = \min(\text{size}_{\text{IP},F_{\text{in}}}(\mathbf{x}), \text{size}_{\text{IP},F_{\text{out}}}(\mathbf{x}))$$

that is, as the size of a sphere that can travel to $\mathbf{x}$ from both the inlet and the outlet. The sixth descriptor, constrictivity, is then defined relative to the sphere size as

$$\beta = M_{\mathbf{x}\in\Gamma}\left(\text{size}_{\text{IP,through}}(\mathbf{x})\right) / M_{\mathbf{x}\in\Gamma}\left(\text{size}_{\text{sph}}(\mathbf{x})\right)$$

with M denoting the median. Note that this is different from the standard constrictivity, which is usually defined with respect to either $\text{size}_{\text{IP},F_{\text{in}}}(\mathbf{x})$ or $\text{size}_{\text{IP},F_{\text{out}}}(\mathbf{x})$ instead of $\text{size}_{\text{IP,through}}(\mathbf{x})$.[14,15] We compute the constrictivity in the pore space, that is, with $\Gamma = \{\mathbf{x} : \mathcal{D}(\mathbf{x}) = 1\}$.

Geodesic tortuosity, the seventh and last descriptor, measures the length of paths through a subset Γ of the window W, and is here computed in the pore space with $\Gamma = \{\mathbf{x} : \mathcal{D}(\mathbf{x}) = 1\}$. For a point $\mathbf{x} \in \Gamma$, the (point-wise) tortuosity $\tau(\mathbf{x})$ equals the relative length of the shortest path through Γ connecting $\mathbf{x}$ to the inlet and outlet faces F_{in} and F_{out} in the chosen direction x. The geodesic distance $\text{geoddist}(\mathbf{x}, S)$ is defined as the length of the shortest path through Γ from $\mathbf{x}$ to the nearest point in a set $S \subset \Gamma$ and the point-wise tortuosity can then be written as

$$\tau(\mathbf{x}) = \frac{\text{geoddist}(\mathbf{x}, F_{\text{in}}) + \text{geoddist}(\mathbf{x}, F_{\text{out}})}{L} \quad (1)$$

where L is the length of the window W in the chosen direction. Geodesic tortuosity is commonly defined as the mean of the point-wise tortuosities at the inlet face,

$$\tau_{\text{in}} = \frac{1}{|F_{\text{in}} \cap \Gamma|} \int_{F_{\text{in}} \cap \Gamma} \tau(\mathbf{x}) d\mathbf{x} \quad (2)$$

The set of tortuosities at the inlet face $F_{\text{in}} \cap \Gamma$ can be computed by the Formula in Equation (1) using only the distance transform $\text{geoddist}(\mathbf{x}, F_{\text{out}})$ as in this case $\text{geoddist}(\mathbf{x}, F_{\text{in}}) = 0$. Thus only $\text{geoddist}(\mathbf{x}, F_{\text{out}})$ is needed to compute $\tau_{\text{in}}(\mathbf{x})$. However, if we compute both distance transforms we obtain the point-wise tortuosities for the full pore space Γ, and can define the full sample geodesic tortuosity

$$\tau = \frac{1}{|\Gamma|} \int_{\Gamma} \tau(\mathbf{x}) d\mathbf{x} \quad (3)$$

Thus we obtain τ at only twice the cost of computing τ_{in}. We prefer to use the full sample τ, as it has been shown to provide substantially better predictive models for transport through pore systems than τ_{in}.[3]

Here, the surface area is computed with the Matlab package `imMinkowski`,[12] the contact distance using the Matlab function `bwdist`,[16] and the 3D sphere size, intrusion porosimetry and geodesic tortuosity with the methods `3D pore size`, `Intrusion porosimetry` and `Geodesic tortuosity`, respectively, in the in-house software Mist.[17]

2.2.2 | Descriptors on experimental data

We compute the descriptors for the experimental data structure shown in Figure 1. First, we compute total porosity ($\epsilon_{\text{total}} = 15\%$) and effective porosity ($\epsilon = 14\%$) which are relatively low. Then, the 3D sphere size is computed both in the solid space (median = 55 voxels) and pore space (median = 18 voxels) and shown in Figure 1B and Figure 1C, respectively. The 3D sphere size is much larger in the solid space due to the volume fraction being higher and to the presence of large calcium carbonate particles. The median of the computed through-intrusion porosimetry sizes in the pore space, illustrated in Figure 1D, is low, 5.0, compared to the 3D sphere size, resulting in a low constrictivity ($\beta = 0.28$). The tortuosity ($\tau = 1.4$) shown in Figure 1E has a relatively inhomogeneous distribution, with low values on the left and right sides of the structure, but relatively high values in the middle (high values are transparent in the visualization).

2.3 | Model

The model for the structure $\mathcal{D}$ can be split into two components, one for large-scale and one for small-scale variation in the coating. The large-scale component corresponds roughly to large particles, and the small-scale component corresponds roughly to small particles, binder, and pore phase. Note, however, that this is not an explicit model for example the large particles, but the large-scale and small-

scale components are tools which interact in encoding the microstructural features at different scales in the data.

The first model component, denoted X_{tess}, corresponds to large-scale variations and determines the location of large voids. It is created using the following steps: (1) a sphere packing; (2) a Laguerre tessellation based on the sphere packing; (3) a skeletonization of the tessellation; and (4) a smoothing of the tessellation using a Gaussian filter with standard deviation σ. Each sphere in the packing is assigned a radius sampled from a log-normal distribution, with coefficient of variation cv_{sph}. Spheres are positioned sequentially in order of descending radius, and for each sphere, a center location where the sphere does not overlap with already placed spheres is suggested by drawing from a uniform distribution in the sample window W. A maximum of 1000 attempts are made, after which a new radius is drawn. This process is continued until a pre-determined number of spheres, n_{sph}, have been placed. The chosen radii are then re-scaled to obtain the desired volume fraction of the sphere packing. A Laguerre tessellation[18] based on weighted distances to define the cells of the tessellation taking into account both the location and radius of each sphere is then computed. The skeleton of the tessellation, encoded as a binary field that equals one on the skeleton and zero elsewhere, is then smoothed with a Gaussian filter with standard deviation σ. The

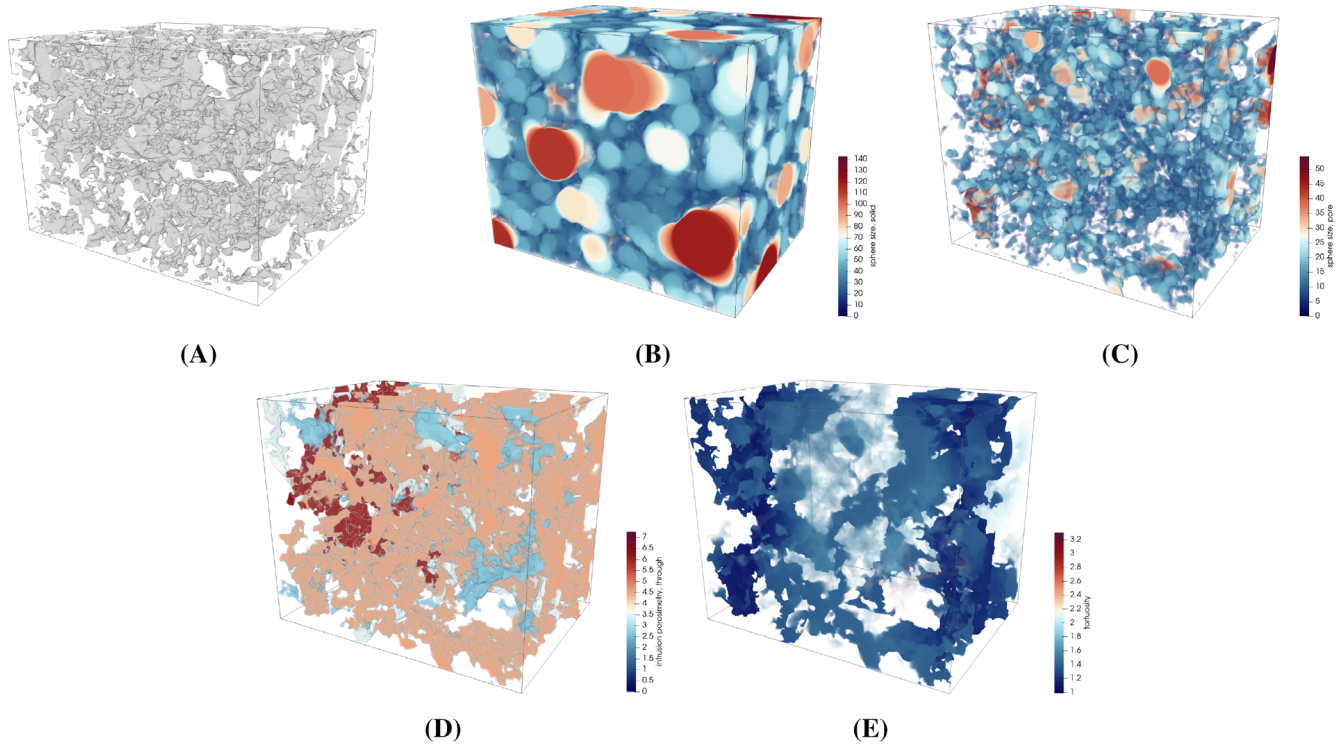

FIGURE 1 (A) Data structure $\mathcal{D}$; the descriptors sphere size in (B) the solid phase and in (C) the pore space; (D) through-intrusion porosimetry; (E) geodesic tortuosity. The transparency for the tortuosity is chosen such that voxels with tortuosities higher than 1.5 are barely visible.

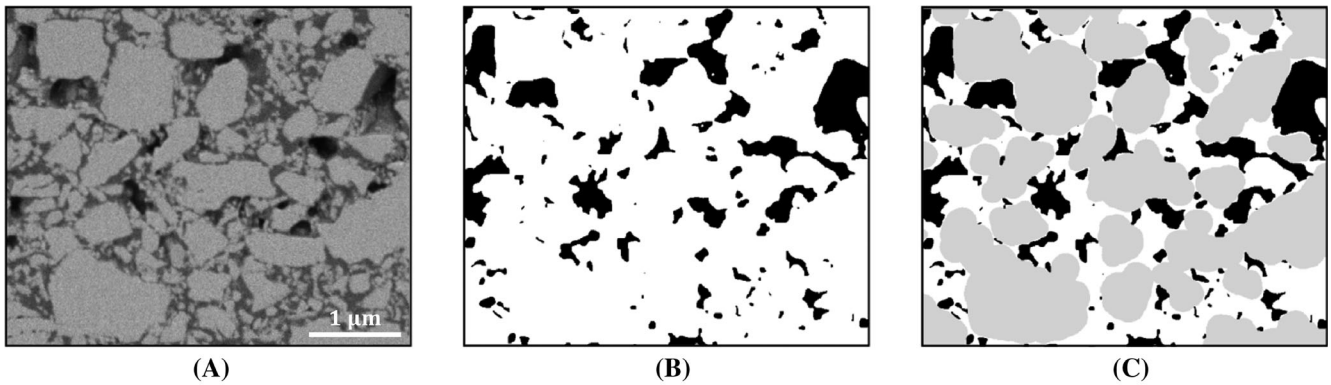

FIGURE 2 (A) FIB-SEM image of the coating. (B) The corresponding segmented image with solid (white) and pore (black) phases. (C) The segmented image with the approximation of large particles overlaid in gray. The black and white areas together constitute the set Γ_{GRF} where the covariance is fitted.

tessellation component X_{tess} is defined as the negative of this smoothed skeleton.

The second model component X_{GRF}, which captures smaller scale variations in the coating, is a stationary and isotropic zero-mean Gaussian random field (GRF).[19,20] We construct the covariance of the GRF using the popular stationary and isotropic Matérn covariance.[21]

$$C(\mathbf{x}_1, \mathbf{x}_2) = c(\kappa \| \mathbf{x}_1 - \mathbf{x}_2 \|)^{\rho} K_{\rho}(\kappa \| \mathbf{x}_1 - \mathbf{x}_2 \|) \tag{4}$$

Here, $c > 0$ is a scaling parameter; $\kappa > 0$ controls the range of the correlation; $\rho > 0$ controls the smoothness of the GRF; and K_{ρ} is the modified Bessel function of the second kind and order ρ.

To define the covariance of X_{GRF}, we use a linear combination of three Matérn covariances on this domain, namely

$$C^{X_{\text{GRF}}} = C_{\text{short}} + C_{\text{mid}} + C_{\text{long}} \tag{5}$$

with separate parameters ρ and κ for each of the covariances C_{short}, C_{mid} and C_{long}. The range of the correlation, controlled by κ, is such that C_{short} has the shortest range and C_{long} has the longest range. The scaling parameters c are chosen such that $C_{\text{short}}(0) = 1$, and $C_{\text{mid}}(0) = \alpha_{\text{mid}}$ and $C_{\text{long}}(0) = \alpha_{\text{long}}$, where the parameters α_{mid} and α_{long} are to be estimated. The final covariance $C^{X_{\text{GRF}}}$ is normalized so that $C^{X_{\text{GRF}}}(0) = 1$.

The two model components, X_{tess} and X_{GRF}, are normalized to lie within the interval $[-1, 0]$ and $[-1, 1]$, respectively. The weighted sum $wX_{\text{tess}} + X_{\text{GRF}}$ of the two model components, where $w > 0$ is a scaling parameter, is then thresholded to create the full model

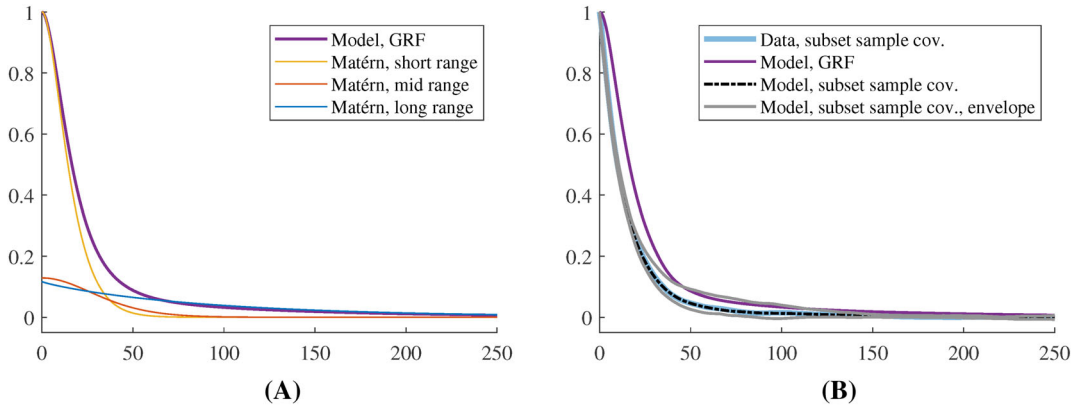

FIGURE 3 (A) Model GRF, with the three Matérn covariances used to create the model GRF. (B) Covariance estimated from the data, showing the subset sample covariance of the data (blue), the fitted model GRF (purple), and the average (black, dashed) and envelope (black, solid) of subset sample covariances computed from 20 samples generated from the GRF and overlaid on the set Γ_{GRF} of the data.

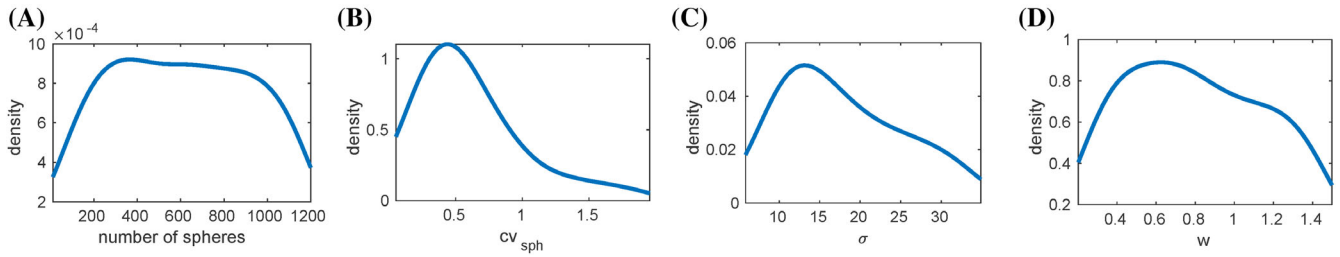

FIGURE 4 Posterior distribution of the tessellation parameters: (A) number of spheres; (B) coefficient of variation cv_{sph}; (C) smoothing σ; and (D) weight w.

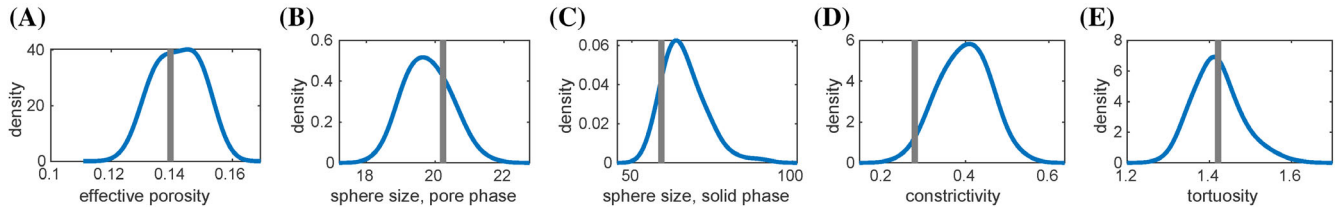

FIGURE 5 Kernel estimates of the density for each of the descriptors effective porosity, 3D sphere size in the pore phase and in the solid phase, constrictivity and tortuosity, computed from 300 model generated structures. The value of the descriptor computed in the data indicated with a vertical black line.

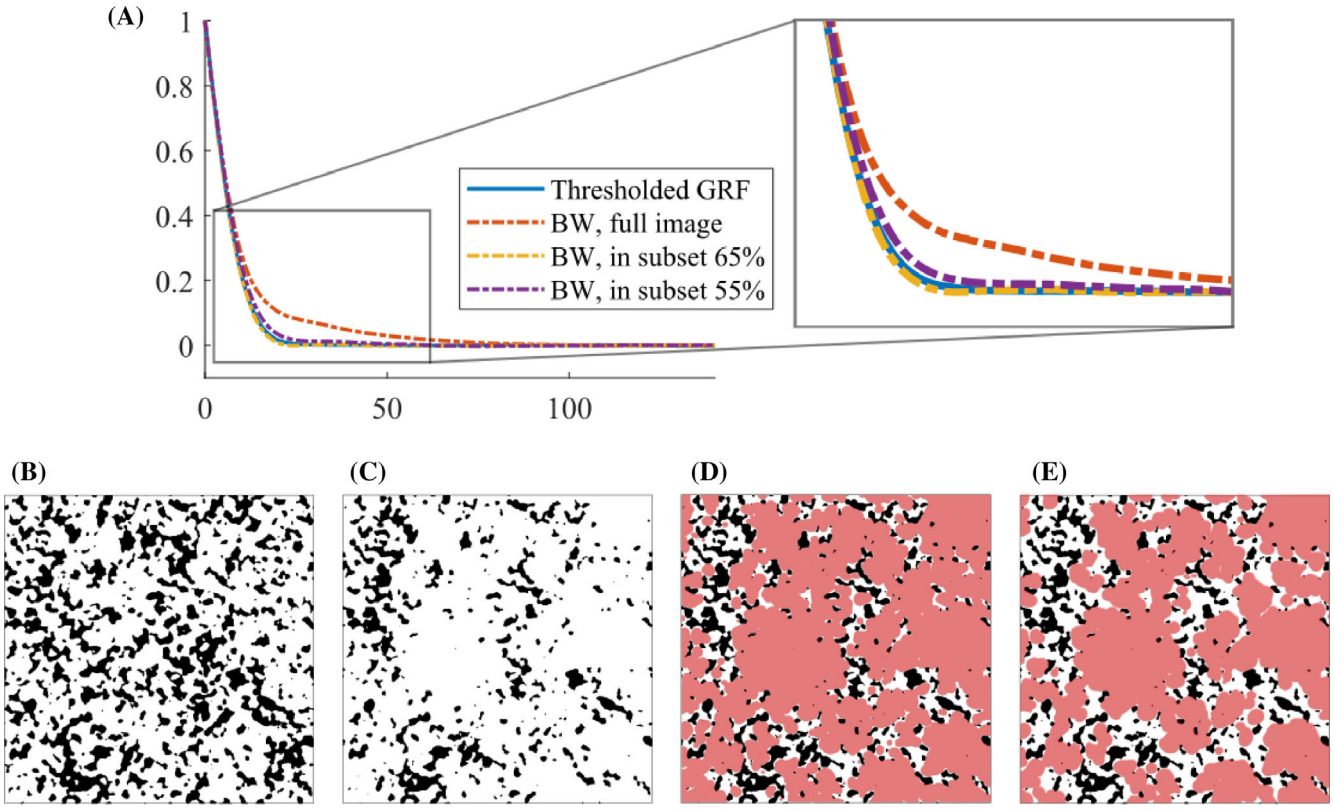

FIGURE 6 Validation of the GRF model fitting. (A) Covariances fitted to a model generated structure $\mathcal{D}$: The full sample covariance $\widehat{C}^{\mathcal{D}}$ (dashed red), and the subset sample covariance $\widehat{C}^{\mathcal{D}}_{\Gamma_{\text{GRF}}}$, with the subset Γ_{GRF} covering either 35% (dashed yellow) or 45% (dashed purple) of W. The target full sample covariance $\widehat{C}^{X_{\text{GRF}} \geq u}$ of the thresholded GRF is shown in solid blue. (B)–(E) 2D slices of the structures used to fit the covariances: (B) the thresholded GRF, (C) the model generated structure $\mathcal{D}$, and (D) and (E) showing the structure $\mathcal{D}$ with the set $\{t : t \notin \Gamma_{GRF}\}$ in red for the subset Γ_{GRF} covering (D) 35% and (E) 45% of W, respectively.

TABLE 1 Validation with regard to the tessellation component X_{tess} of the model, showing the true value of the parameters used to generate the model structure $\mathcal{D}$; the range of values of each parameter in the ABC; and the posterior medians of the ABC sample.

	n_{sph}	cv_{sph}	σ	w
True values	500	0.6	20	1
ABC range	[10, 1200]	[0.05, 2]	[5, 35]	[0.2, 1.5]
Parameter estimate	570	0.56	22	0.96

$$\mathcal{D}(\mathbf{x}) = I_{wX_{\text{tess}} + X_{\text{GRF}} \geq u}(\mathbf{x}) \tag{6}$$

where $I_{\text{condition}}(\mathbf{x})$ is the indicator function which equals one if condition holds at location $\mathbf{x}$ and zero otherwise. The threshold u is selected so that the effective porosity of $\mathcal{D}$ matches the effective porosity of the data.

X_{tess} takes high values in the cells of the tessellation, and low values on and around the skeleton. In regions with high values of X_{tess}, in the interior of tessellation cells, the sum in Equation (6) is dominated by X_{tess} (assuming that the scaling parameter w is large enough) and so these regions will be part of the solid phase ($\mathcal{D}(\mathbf{x}) = 0$). Regions within large tessellation cells correspond roughly to the gray areas in Figure 2C (for more details related to Figure 2, see Section 2.4). X_{GRF} on the other hand determines the local structure of $\mathcal{D}$ in areas on and around the (smoothed) skeleton corresponding to the black and white areas in Figure 2(C).

We enforce periodicity for the model components X_{tess} and X_{GRF}, since this removes the problem of boundary effects at the edges of the observation window when computing microstructural descriptors and permeability.

In practice, we work with discrete (lattice) realizations of the fields X_{tess} and X_{GRF}. For the GRF, we use the convenient block-circulant form of the covariance matrix of the GRF observed on the lattice, and generate realizations of X_{GRF} using the computationally efficient discrete Fourier transform, see for example, [22, Chapter 2.6] and [23, Chapter 12.2.2].

2.4 | Model fitting

2.4.1 | Model fitting approach

We present a two step approach to fit the proposed model to the data, that is, the microstructure of the coating. We first fit the X_{GRF} model component by using a novel method which estimates the covariance in a subset of the microstructure, where the subset is

identified using the microstructural descriptor 3D sphere size. Once the X_{GRF} model component is fixed, the parameters related to the X_{tess} model component and the scaling w used when combining the components are estimated using approximate Bayesian computation (ABC).

Estimation of X_{GRF} by covariance estimation in a complex geometry

The GRF of the second model component is mainly intended to model the structure in between large particles, and so we suggest to fit the covariance $C^{X_{GRF}}$ of the GRF in this subset. To achieve this, we construct a method that approximates the locations of the large particles and estimates $C^{X_{GRF}}$ semi-parametrically in the complement of this set.

To approximate the locations of the large particles, we used the 3D sphere size in the solid phase of the microstructure. Since the 3D sphere sizes in voxels contained within the large particles are considerably larger than in voxels belonging to the binder, the locations of large particles are approximated by the set of voxels with sphere sizes larger than a threshold u_{size}, with u_{size} determined by visual inspection.

TABLE 2 Ranges for the parameters in the flow simulation study.

	Parameter range
ϵ	[15%, 30%]
α_{mid}	[0.05, 20]
α_{long}	[0.05, 20]
cv_{sph}	[0.3, 1.5]
σ	[5, 30]
w	[0.3, 3.5]

The GRF X_{GRF} is then fitted to the complement of this set, $\Gamma_{GRF} = \{\mathbf{x} : size_{sph}(\mathbf{x}) < u_{size}\}$, which we assume consists mainly of binder, pores and smaller particles, see the black and white areas in Figure 2C.

The sample covariance estimated over a subset $\Gamma \subseteq W$ can be written as

$$\widehat{C}_{\Gamma}(v) = \frac{1}{n} \sum_{\mathbf{x} \in \Gamma : \mathbf{x}+v \in \Gamma} \mathcal{D}(\mathbf{x})\mathcal{D}(\mathbf{x}+v), \quad v \in W \tag{7}$$

where n is the size of the set $\{\mathbf{x} \in \Gamma : \mathbf{x}+v \in \Gamma\}$. Normally, the set Γ would be the full observation window W, in which case we use the term *full sample covariance*. The formula, however, works just as well when Γ is a subset of W, Γ_{GRF} in our case, resulting in the *subset sample covariance* $\widehat{C}_{\Gamma}$.

Note that X_{GRF} will actively influence the structure $\mathcal{D}$ only where the tessellation based component X_{tess} takes relatively low values compared to X_{GRF}, which will mainly be on and around the tessellation skeleton. The set Γ_{GRF} is designed to capture these regions where the GRF is "active" (corresponding to regions in black and white in Figure 2(C)). To fit only the GRF component, we make the simplifying assumption that X_{tess} only influences the structure through the set Γ_{GRF}. More precisely, we estimate the covariance of X_{GRF} for the simplified model $\mathcal{D}' = I_{X'_{tess}+X_{GRF} \geq u}$, where $X'_{tess}(\mathbf{x})$ equals zero for $\mathbf{x} \in \Gamma_{GRF}$ and u otherwise. Within model D', when we restrict the model to the set Γ_{GRF}, the model is equivalent to the thresholded GRF $I_{X_{GRF} \geq u}$. Thus, to fit $\mathcal{D}'$ to the data, we need as the first step the set Γ_{GRF} computed in the data, and then, to find the best fit for X_{GRF} given the assumption that the data is a realization of $I_{X_{GRF} \geq u}$ within this set.

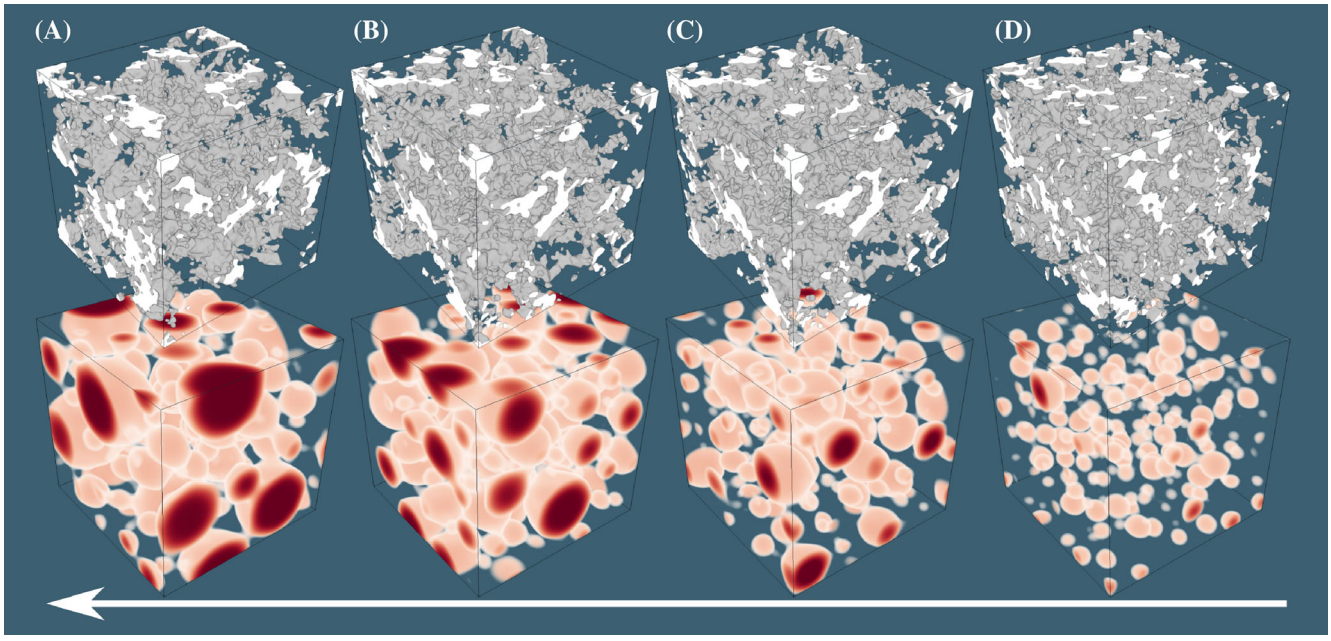

FIGURE 7 Four structures generated with values of the coefficient of variation parameter cv_{sph} (A) 1, (B) 0.75, (C) 0.5, and (D) 0.3. The arrow indicates increasing values of the parameter. The top row shows the structure $\mathcal{D}$ and the bottom row shows a thresholded version of the tessellation component X_{tess} used to create $\mathcal{D}$.

An estimate of $C^{X_{GRF} \geq u}$ within Γ_{GRF} is obtained by computing the subset sample covariance $\widehat{C}^{data}_{\Gamma_{GRF}}$ in the data restricted to Γ_{GRF}. Then, to obtain an estimate of the covariance $C^{X_{GRF}}$ of the non-thresholded GRF, we use the following one-to-one relationship

$$C^{X_{GRF} \geq u}(\mathbf{x}) = \int_0^{C^{X_{GRF}}(\mathbf{x})} \varphi(u, u; z)\mathrm{d}z \tag{8}$$

which holds when the GRF is stationary with zero-mean and unit variance, see references [13,24: pp. 26–27]. Here $\varphi(\cdot, \cdot; z)$ is the density of a bivariate Gaussian distribution with correlation z. We choose the threshold u so that $I_{X_{GRF} \geq u}$ has the same effective porosity as the data restricted to the set Γ_{GRF}. Replacing $C^{X_{GRF} \geq u}$ by $\widehat{C}^{data}_{\Gamma_{GRF}}$ and solving numerically for $C^{X_{GRF}}$ in Equation (8), we then get an estimate $\overline{C}_{\Gamma_{GRF}}$ of the covariance $C^{X_{GRF}}$ of X_{GRF}. After this, the estimate $\overline{C}_{\Gamma_{GRF}}$ is further fine-tuned so that the sample covariances of the subsets $\widehat{C}^{X_{GRF} \geq u}_{\Gamma_{GRF}}$ and $\widehat{C}^{data}_{\Gamma_{GRF}}$ are as close as possible to each other.

Finally, a linear combination of Matérn covariances that approximates $\overline{C}_{\Gamma_{GRF}}$ well was fit, resulting in the following parameter estimates for $C^{X_{GRF}}$: $\widehat{\kappa} = 0.14$ and $\widehat{\rho} = 1.9$ for the short range C_{short}, $\widehat{\kappa} = 0.11$ and $\widehat{\rho} = 5.3$ for the mid range C_{mid}, $\widehat{\kappa} = 0.01$ and $\widehat{\rho} = 0.4$ for the long range C_{long}, and $\widehat{\alpha}_{mid} = 0.13$ and $\widehat{\alpha}_{long} = 0.12$ for the scaling parameters used to create $C^{X_{GRF}}$. All four covariances, C_{short}, C_{mid}, C_{long}, and $C^{X_{GRF}}$, are shown in Figure 3(A). Note that since for our data, the subset sample covariance $\widehat{C}^{data}_{\Gamma_{GRF}}$ can be considered isotropic, we compute one-dimensional sample covariances by averaging 3D sample covariances over spherical shells.

Figure 3B shows a comparison between the subset sample covariance $\widehat{C}^{data}_{\Gamma_{GRF}}$ computed from the data, and the average $\overline{C}_{\Gamma_{GRF}}$ of the 20 subset sample covariances from simulations of the thresholded GRF.

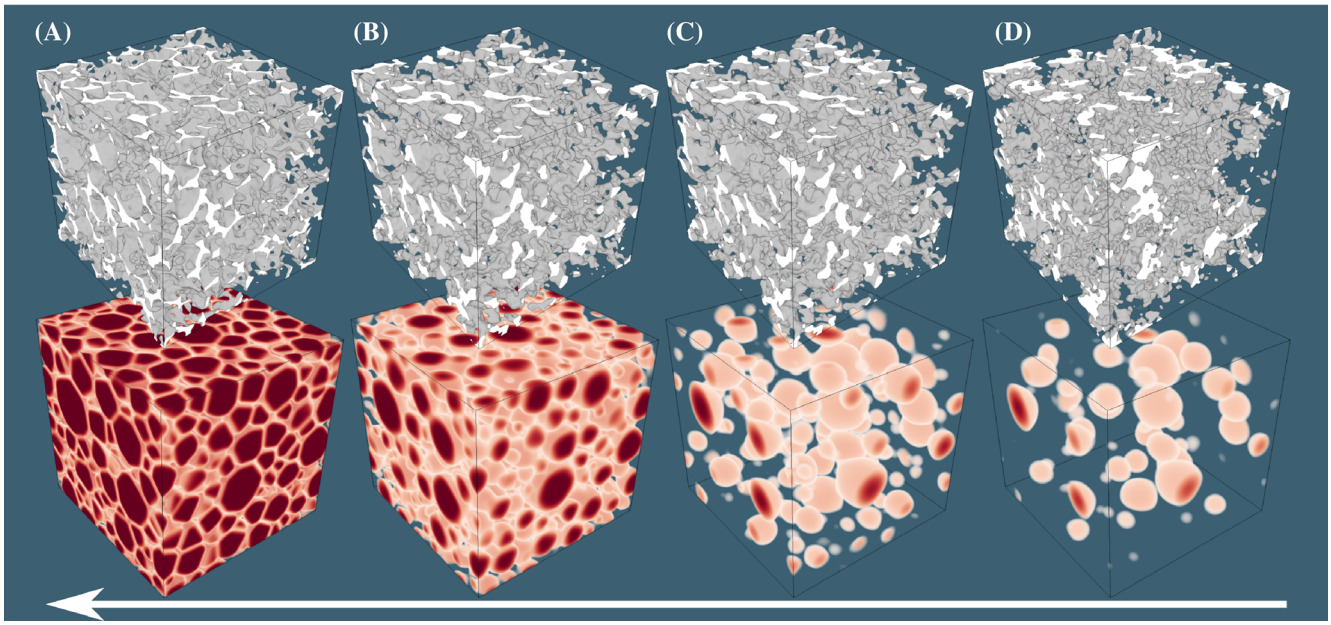

FIGURE 8 Four structures generated with values of the smoothing parameter σ (A) 5, (B) 10, (C) 15, and (D) 20. The arrow indicates decreasing smoothing. The top row shows the structure $\mathcal{D}$ and the bottom row shows a thresholded version of the tessellation component X_{tess} used to create $\mathcal{D}$.

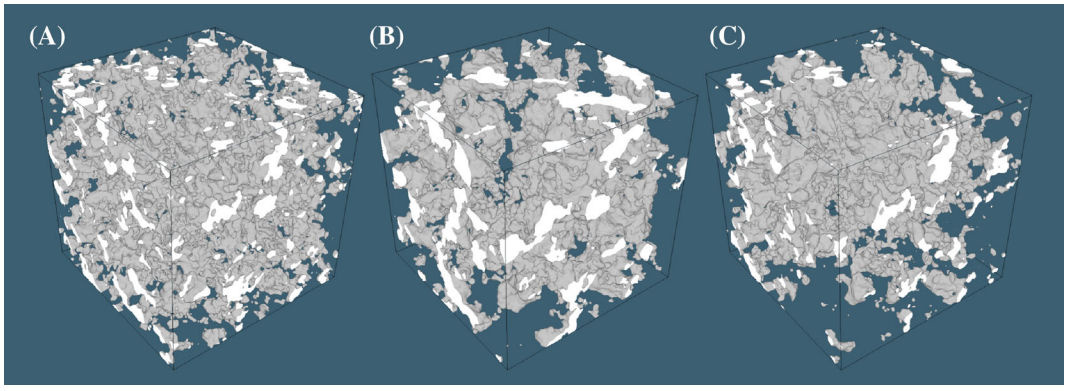

FIGURE 9 Examples of structures with varying values of the parameters $(\alpha_{mid}, \alpha_{long})$, with (A) (1, 1), (B) (10, 1), and (C) (1, 10).

The averages over the 20 GRF realizations (black, dashed) are very close to the data (blue).

Estimation of X_{tess} by approximate Bayesian computation

Due to the complexity of the model, we turn to approximate Bayesian computation (ABC), a method for simulation-based posterior inference, for parameter estimation in the tessellation component X_{tess}. Classically, we would write the posterior distribution as

$$f(\boldsymbol{\theta}|\mathcal{D}) \propto P(\mathcal{D}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}), \tag{9}$$

where $\boldsymbol{\theta}$ is the parameter vector, $\mathcal{D}$ is the data, $P(\mathcal{D}|\boldsymbol{\theta})$ is the likelihood, and $\pi(\boldsymbol{\theta})$ is the prior distribution. In our case, the likelihood function is intractable and ABC provides a way of sampling from an approximate posterior distribution.[25–27] If we sample $\boldsymbol{\theta}$ from $\pi(\boldsymbol{\theta})$, simulate a structure $\mathcal{D}'$ and accept $\boldsymbol{\theta}$ if $\mathcal{D}' = \mathcal{D}$ and reject otherwise, we are sampling from the exact posterior. However, we can relax this condition by only demanding that $\boldsymbol{\rho}(\mathcal{D}, \mathcal{D}') \leq \epsilon$ for some discrepancy measure(s) $\boldsymbol{\rho}$ and tolerance(s) ϵ, thereby performing approximate but computationally much more efficient inference. The usual posterior inference, that is, computing the mean and credible intervals, can then be performed for the approximate posterior. Here, we use basic accept-reject ABC although more sophisticated (e.g., Markov chain-based and population-based) approaches exist.[28–32] As discrepancy measures, we use the microstructural descriptors recalled in Section 2.2.1. One possibility would be to incorporate a set of different microstructural descriptors into a single (scalar) discrepancy measure, but here it turned out more practical to define different tolerances for different descriptors. Specifically, we perform the accept-reject step in the following manner. We consider as discrepancy measures the relative errors of the descriptors,

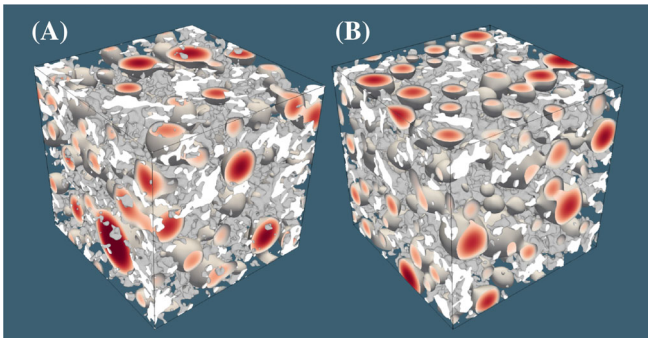

FIGURE 10 Two structures with tessellation weight w (A) 0.5 and (B) 2.5, respectively.

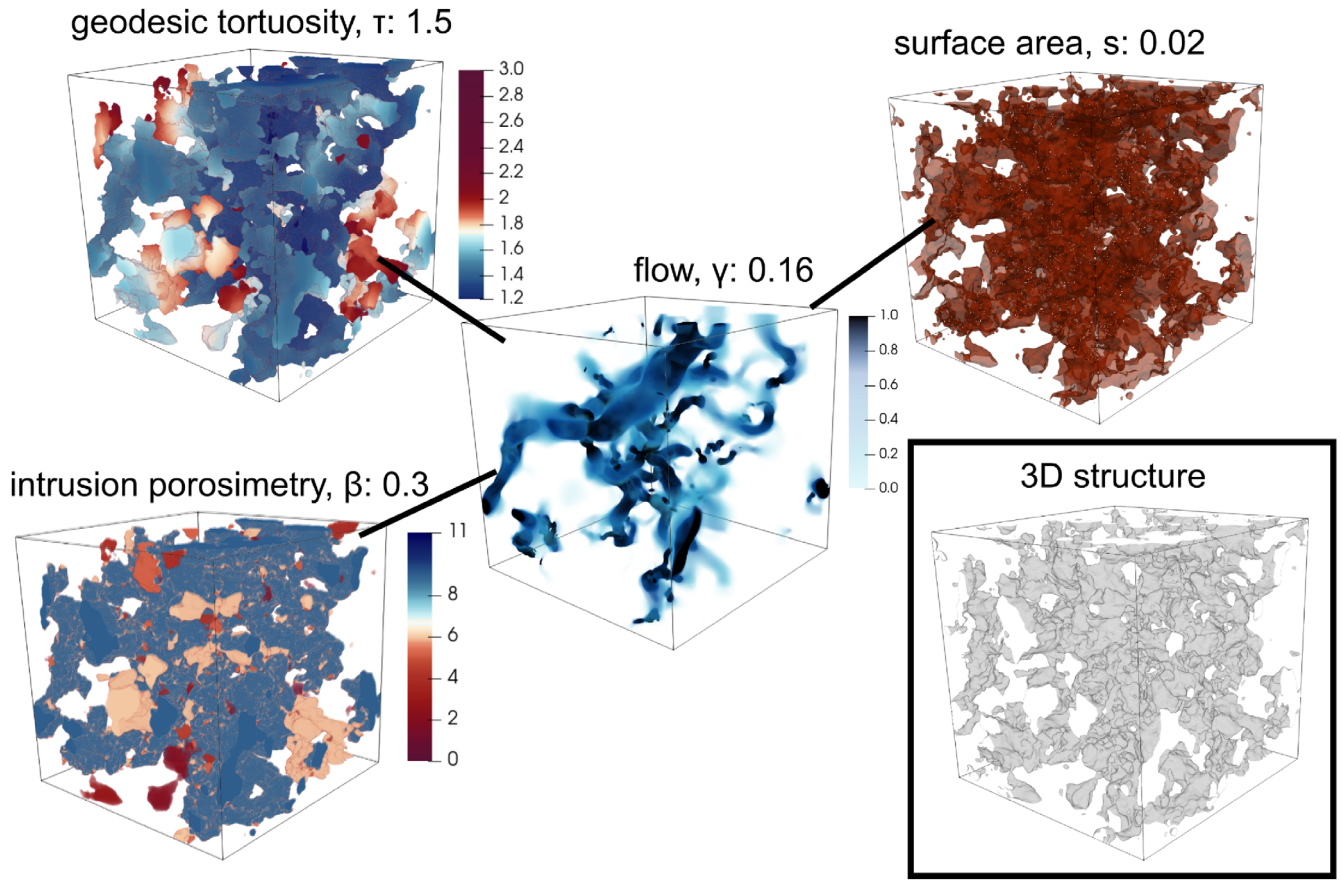

FIGURE 11 One structure from the flow simulation study (bottom right) together with the numerically simulated flow field (middle), geodesic tortuosity (top left), surface between pore and solid phase (top right), and intrusion porosimetry (bottom left) within the structure.

$$\rho_{\text{descriptor}} = \frac{|\,\text{descriptor}_{\text{model}} - \text{descriptor}_{\text{data}}\,|}{\text{descriptor}_{\text{data}}}. \quad (10)$$

First, we choose to accept a sample only if $\rho_{\text{surface area}} \le 0.25$. Second, we accept a sample only if $\rho_{\text{descriptor}}$ is less than the 75%-quantile of $\rho_{\text{descriptor}}$, where the quantile is computed over the distribution of the approximate ABC posterior based on the samples where $\rho_{\text{surface area}} \le 0.25$. As descriptors, we use surface area, mean contact distance in pore phase and in solid phase, mean of tortuosity and mean and standard deviation of sphere size in the solid phase. The first step resulted in approximately 2000 model structures and the second step in approximately 1400 structures. Given that some of the posterior distributions shown in Figure 4 are heavy tailed, the posterior medians of the ABC sample (i.e., the sample from the approximate posterior ABC distribution) were chosen as point estimates for the parameters resulting in $\hat{n}_{\text{sph}} = 600$, $\hat{\text{cv}}_{\text{sph}} = 0.5$, $\hat{\sigma} = 16$, and $\hat{w} = 0.75$.

To assess the goodness-of-fit of the model in Equation (6), we use the descriptors effective porosity; covariance; 3D sphere size in the solid and pore phase; geodesic tortuosity; and constrictivity. Figure 5 compares the values of descriptors computed in 300 structures generated from the fitted model, with the value of the descriptor computed in the data. The model fit is good for all descriptors except constrictivity, as can be seen by comparing the histogram of descriptor values computed in structures generated from the model with those computed from the data (Figure 5). Here the constrictivity value observed in the data is in the lower range of simulated values from the model.

There is a trade-off between the model fit regarding the covariance and the sphere size. If model parameters are changed so that the relative model fit with regard to the covariance goes up, then the relative fit with regard to the 3D sphere size in the solid phase goes down, and vice versa. Thus the model fit cannot be good for both descriptors simultaneously.

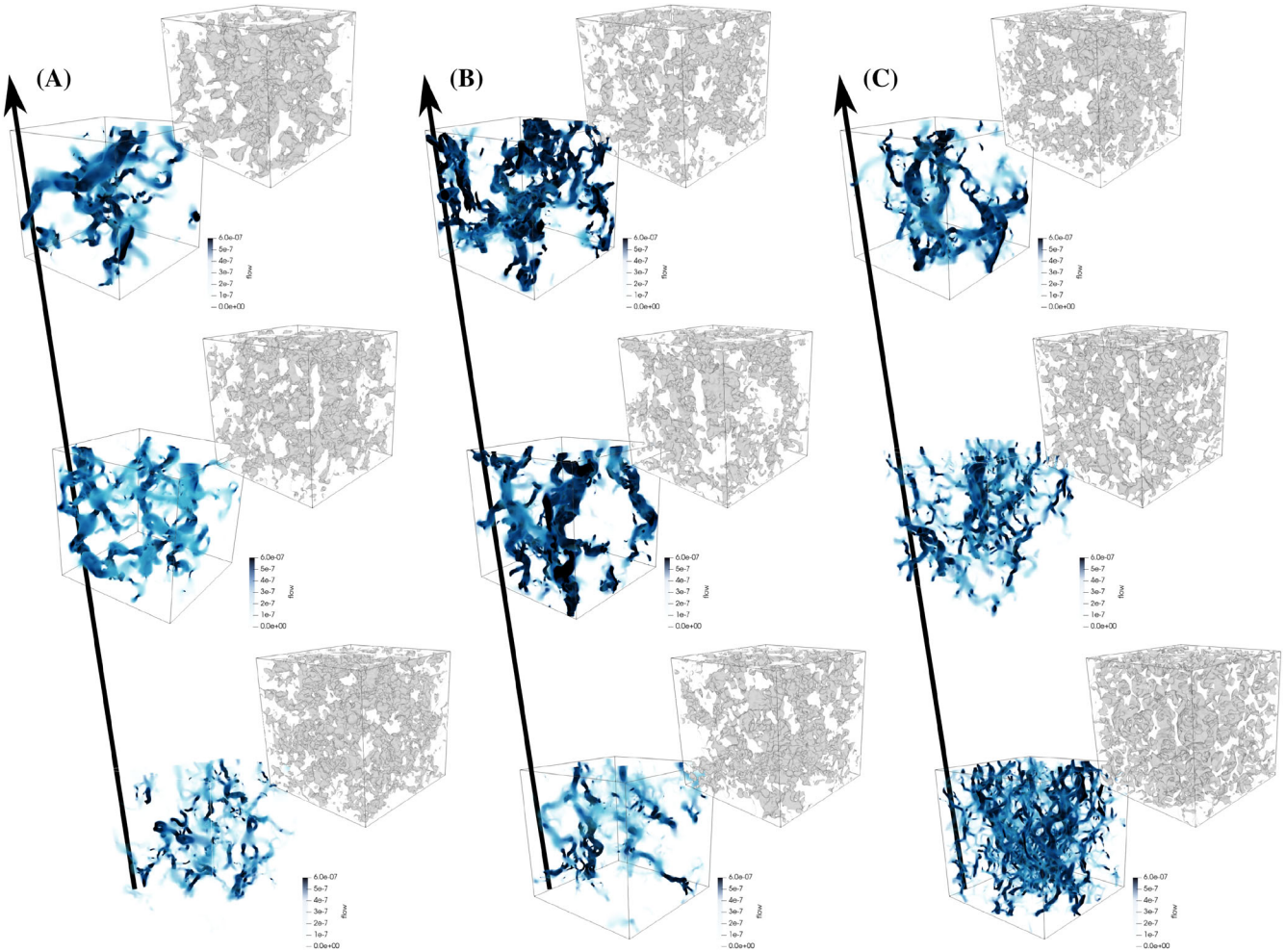

FIGURE 12 Nine of the 800 structures from the flow simulation study, showing the numerically simulated flow in 3D. For each structure, the permeability γ computed from the 3D flow, and the microstructural descriptors mean geodesic tortuosity τ, constrictivity β, specific surface area s, and effective porosity ϵ are shown. The nine structures have similar effective porosity values and are ordered (from the top left to the bottom right) by decreasing permeability. The color scale, which is the same across the structures, shows the magnitude of the simulated flow vector in voxels in the pore space. The color scale has been capped at a value of around 10% of the maximum flow and then normalized to the unit interval. The opacity is low for voxels with a small flow magnitude.

2.4.2 | Assessment of the model fitting approach

To assess the performance of the model fitting approach, we conducted a small study, where the approach was used to estimate the parameters from structures $\mathcal{D}$ generated from the model.

The GRF model component

We constructed structures using the model parameters estimated from the real data structure, except that we replaced the estimated GRF by a GRF with a relatively short correlation range. The latter choice was made to make the difference between estimating with subset sample covariance and full sample covariance more clear. Four structures shown in Figure 6 were included, namely (b) the GRF thresholded so that the structure has effective porosity 14%; (c) the model structure $\mathcal{D}$, constructed with the same GRF as in (b); and $\mathcal{D}$ with the subset Γ_{GRF} covering either (d) 35% or (e) 45% of the window W.

In Figure 6A, covariances fitted to the four model structures are shown. Unlike the covariance fitted to the full image, the subset sample covariances $\widehat{C}^{\mathcal{D}}_{\Gamma_{GRF}}$ (dashed yellow and purple) provide estimates that are close to the target $\widehat{C}^{X_{GRF} \geq u}$ (solid blue) showing that the estimation approach works well. Both the 35% and the 45% subsets result in estimated covariances that are close to the target although the 35% gives a slightly better fit. We want to emphasize that the subset sample covariances are considerably closer to the target than what the full sample covariance $\widehat{C}^{\mathcal{D}}$ (dashed red) is, which illustrates the necessity of restricting to relevant subsets when estimating the covariance.

We can conclude that the method is not overly sensitive to either the size of the subset Γ_{GRF}, nor to the simplification made in model $\mathcal{D}'$ compared to $\mathcal{D}$ when fitting the GRF.

The tessellation component

Assessing the ABC procedure to estimate the tessellation component parameters, we investigated which microstructural descriptors to use as discrepancy measures by applying the procedure to a structure $\mathcal{D}$ from the model generated with the parameters estimated from the data. One important consideration is to use descriptors that are relatively quick to compute, as we obtain better estimates with larger posterior ABC samples. Another consideration is to make sure to use descriptors that capture sufficient relevant features of the geometry. Here, we used the microstructural descriptors surface area, contact distance in the pore space, contact distance in the solid phase, geodesic tortuosity, and 3D sphere size of the solid phase, i.e. the same setup as the one applied to the data. In the first iteration, the 3D sphere size was not included as it is relatively time-consuming to compute. However, we found that two of the parameters, namely cv_{sph} and w, were

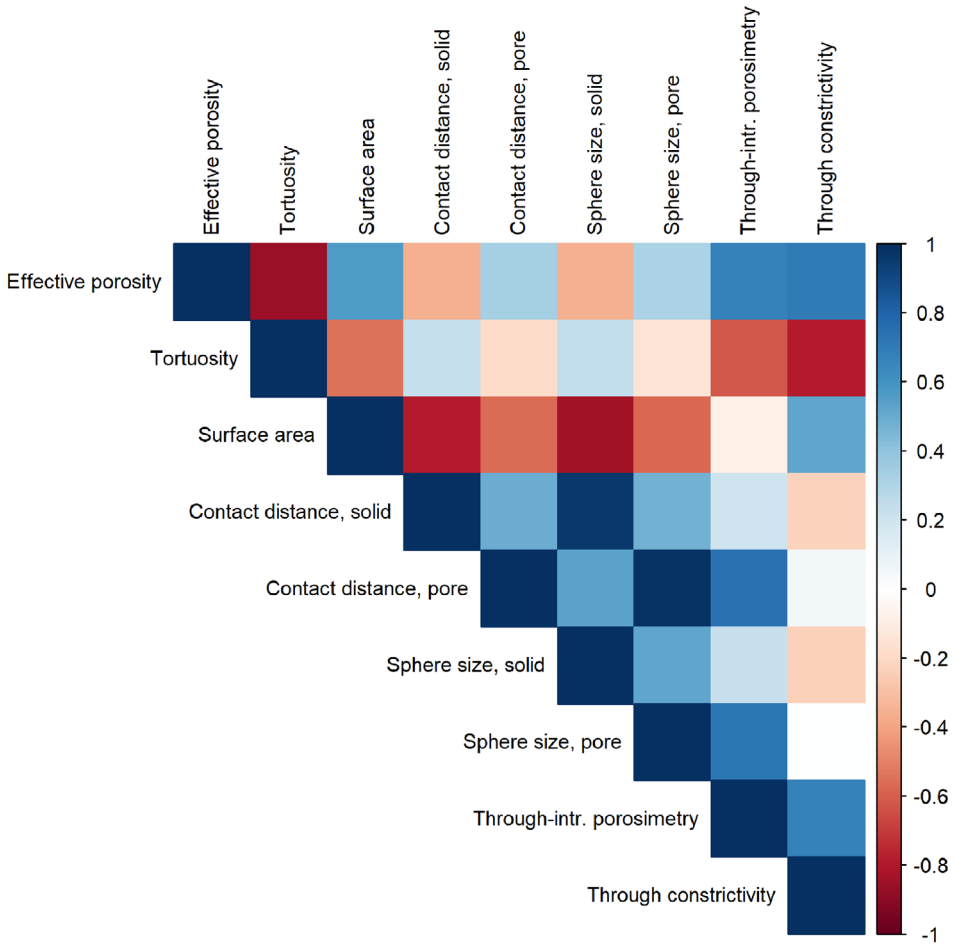

FIGURE 13 Correlation matrix for all microstructural descriptors.

considerably better estimated when we added the 3D sphere size. The final parameter estimates are given in Table 1.

2.5 | Flow simulation study

The variety of structures that can be generated from the model in Equation (6) is explored in a flow simulation study. The set of structures was chosen by varying the model parameters in such a way that the generated structures cover a wide range of microstructural descriptors. First, parameters estimated from the data were used. Then, together with porosity, five model parameters were varied one at a time, keeping all other parameters constant. This created five sets of structures, with 880 structures $\mathcal{D}$ in total.

Besides the tessellation component parameters cv_{sph}, σ and w, two parameters α_{mid} and α_{long} controlling the ratio of the short range (C_{short}), medium range (C_{mid}), and long range component (C_{long}) of the covariance function were included. The transformed covariance $C^* = C_{short} + \alpha_{mid}C_{mid} + \alpha_{long}C_{long}$ was then used to generate the GRF component X_{GRF}. The ranges chosen for the five parameters are shown in Table 2. The range for each parameter was chosen so as to get a wide range of structures, while at the same time staying sufficiently similar to the data to be relevant for permeability comparison. For each set of parameter values, the porosity was varied between 15% and 30%. Parameter values were chosen randomly using a uniform distribution within the chosen intervals. The corresponding structure groups are termed $D_{cv_{sph}}$, D_σ, D_w, $D_{GRF,mid}$, $D_{GRF,long}$. Figures 7–10 show examples of geometries from each of these five groups included in the flow simulation study.

Flow through each structure $\mathcal{D}$ was computed by solving the Navier–Stokes equations within the pore space using the lattice Boltzmann method and in-house software. Examples of the lattice Boltzmann method as implemented in the software can be found in, for example, references [33,34] A pressure drop which drives the flow is applied in direction x, and the permeability κ of the structure is computed by using

$$\bar{u}_x = -\frac{\kappa}{\eta L_x}\Delta p \tag{11}$$

where $\bar{u}_x$ is the average flow in direction x, η is the dynamic viscosity of the fluid, L_x is the thickness of the structure in direction x, and Δp is the applied pressure drop. Numerical stability was ensured by choosing a small pressure gradient to drive the flow, resulting in small Reynolds and Mach numbers. The permeability is independent of the fluid provided that the Reynolds number is sufficiently small, which is the case here.

The convergence criteria for the flow simulations is that the ratio $\sigma_{|vel|}/\mu_{|vel|}$ is less than a set tolerance tol, where $\mu_{|vel|}$ and $\sigma_{|vel|}$ are the mean and standard deviation, respectively, of the flow velocity norm over 1000 iterations. Values tol $= 10^{-4}, 10^{-5}$ and 10^{-6} were explored for three samples. Differences in permeability were on the order of 1%–5% comparing results with tol $= 10^{-4}$ and 10^{-6}, and 0.3%–1% comparing results with tol $= 10^{-5}$ and 10^{-6}.

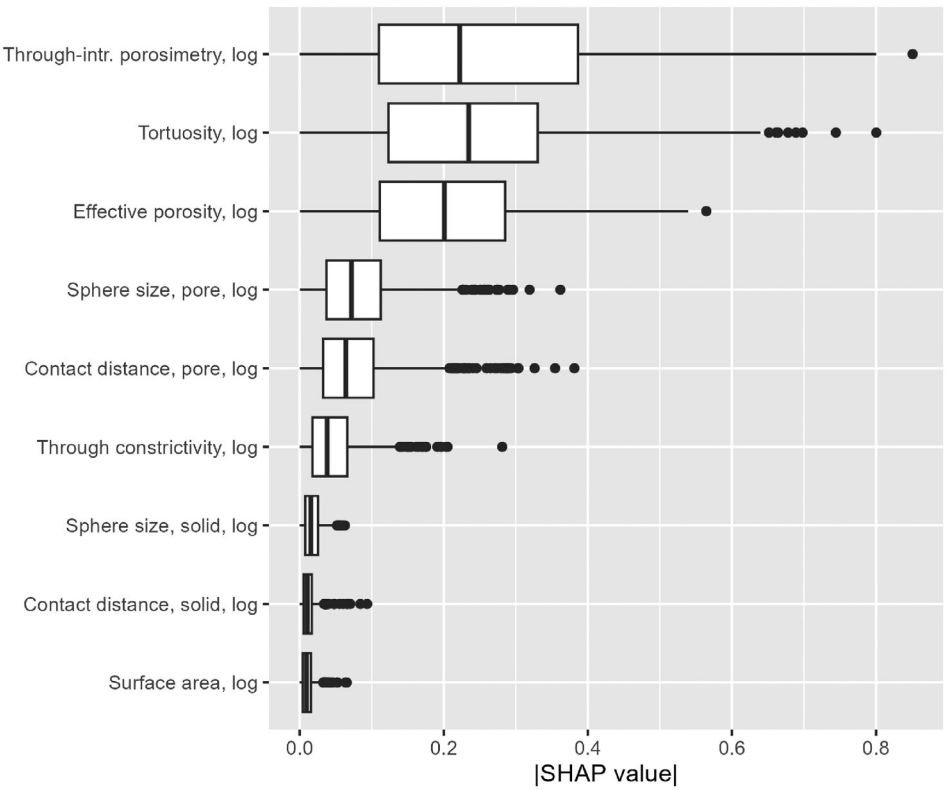

FIGURE 14 Feature importance for permeability with SHAP values. Boxes show the distribution of absolute SHAP values.

The number of iterations required were of the order 10.000 for tol $= 10^{-4}$, 50.000 for tol $= 10^{-5}$ and 500.000 for tol $= 10^{-6}$. The relatively small gain in permeability precision, compared to the long simulation time, when decreasing the tolerance makes a case for using tol $= 10^{-4}$, which is what was used for the permeability study.

The geometry of the structures was characterized by microstructural descriptors, with the purpose of correlating the pore geometry

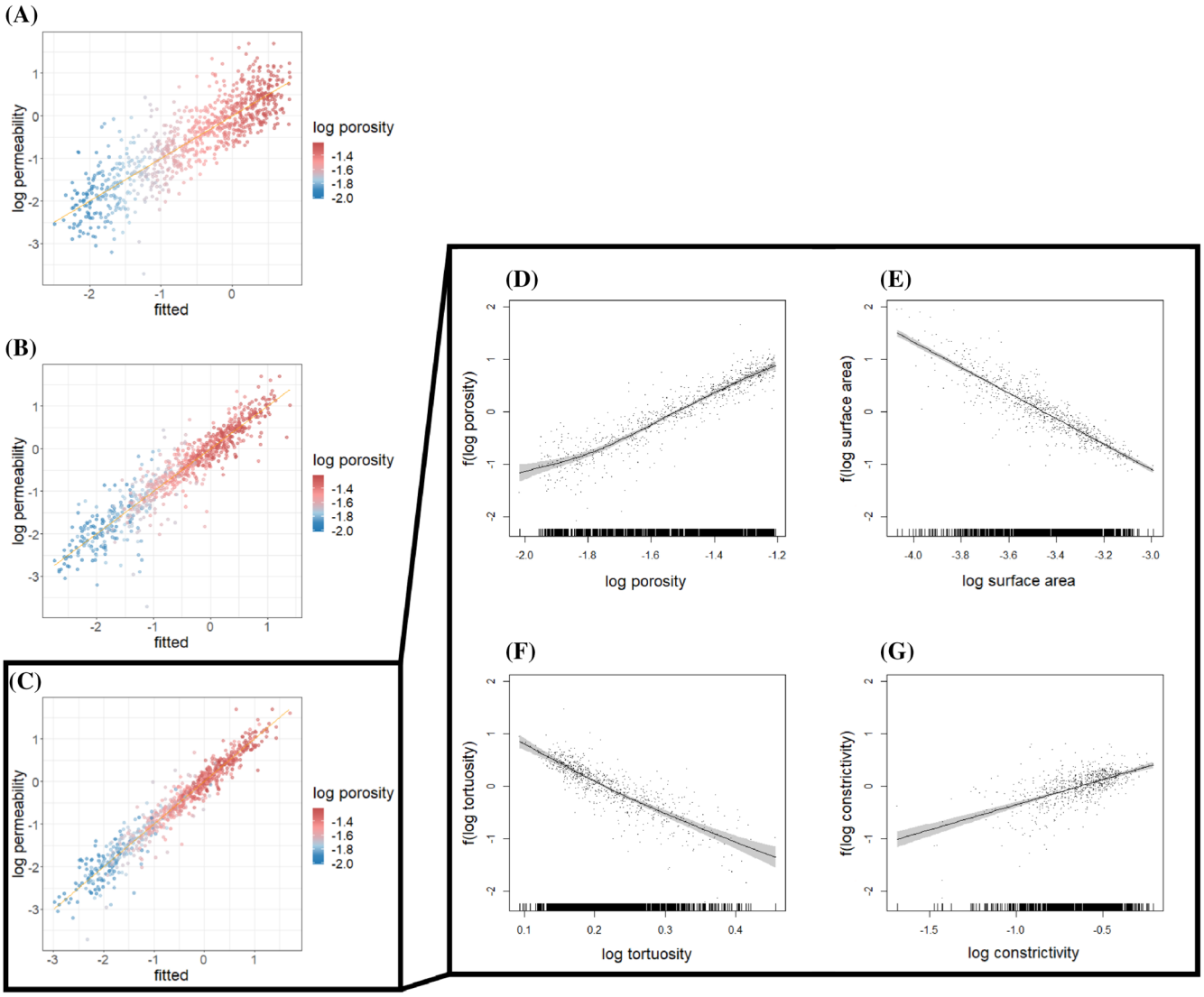

FIGURE 15 Relationship between the response variable log(permeability) and its fitted values, for additive models (12) including predictors (A) effective porosity; (B) effective porosity and specific surface area; and (C) effective porosity, specific surface area, tortuosity and constrictivity as predictors. The fitted smooth function f for each predictor used in the last model is shown in (D)–(G).

TABLE 3 Results for regression models with four sets of predictors, combining effective porosity (ϵ), specific surface area (s), geodesic tortuosity (τ), constrictivity (β), through intrusion porosimetry ($desc_{IP}$) and pore sphere size ($desc_{PS}$).

	GAM		Linear
Predictors	**Full dataset [R^2 (AIC)]**	**Cross-validation [R^2]**	**Full dataset [R^2 (AIC)]**
ϵ	0.73 (1262)	0.72	0.73 (1262)
ϵ, s	0.86 (692)	0.85	0.86 (698)
ϵ, s, τ, β	0.94 (29)	0.93	0.93 (65)
ϵ, τ, $desc_{IP}$, $desc_{PS}$	0.94 (−12)	0.93	0.93 (63)

Note: The four sets of predictors are fitted with the additive model (GAM) and with a linear model (Linear), and results are given either for the full dataset, or for the average for test sets of a 10-fold cross-validation. R^2 values are given for all columns and AIC for models fitted to the full dataset.

of model generated structures with their simulated permeability κ. Additive Gaussian regression models, specifically models of the form

$$y_i = \alpha + \sum_k f_k\left(v_i^{(k)}\right) + e_i, \quad i = 1,\ldots,n, \quad k = 1,\ldots,m \tag{12}$$

where $\mathbf{y} = (y_1,\ldots,y_n)$ is the response variable; α is an intercept; $\mathbf{v}^{(k)} = \left(v_1^{(k)},\ldots,v_n^{(k)}\right)$, $k = 1,\ldots,m$, are predictors, $v_i^{(k)}$ is the value of the kth predictor for the ith data-point; $f_k(\cdot), k = 1,\ldots,m$, are smooth functions; and $\mathbf{e} = (e_1,\ldots,e_n)$ is a vector of i.i.d. mean-zero Gaussian random variables, were fitted. These non-linear regression models are compared with linear models, that is, models where the smooth functions $f_k(\cdot)$ are replaced by linear functions.

We used log-transformed response variable and predictors. Each data-point $\left(v_i^{(1)},\ldots v_i^{(m)}, y_i\right)$ then represents one structure generated from the model in Equation (6), with $s_i^{(k)}$ being the log-transformed value of the microstructural descriptor k computed in that structure, and y_i the log-transformed permeability κ simulated in that structure. The non-linear regression models were fitted using the `gam` function from the R library `mgcv`.[35]

The flow field for one structure is illustrated in Figure 11. Note that the parts of the flow field (middle panel) with high flow coincide with the parts of the structure that have both low geodesic tortuosity (top left panel) and high intrusion porosimetry (bottom left panel). The figure illustrates the range of flow profiles within the dataset, with flow through thin channels (Figure 11B,D) or through few and thick channels (Figure 11A,C); and with well-distributed networks with low tortuosity and high constrictivity (Figure 11A–C) or more poorly distributed networks with high tortuosity and/or low constrictivity (Figure 11G–I). The 3D flow field and the descriptors representing each of the nine structures are shown in Figure 12. The transparency is high for voxels with small flow magnitude, and so only parts of the structure with relatively high flow magnitude can be seen in the 3D visualizations.

Before applying the models, we evaluate the microstructural descriptors for possible redundancy. Figure 13, showing correlations between microstructural descriptors, indicates that there are families of descriptors which provide similar structural information. For example, surface area, contact distance and sphere size, are closely related. The connectivity descriptors tortuosity, through-intrusion porosimetry and constrictivity are also positively correlated. The results indicate that care should be taken when including several correlated predictors into one regression model.

The microstructural descriptors we use as predictors for three of the predictor sets, that is, the effective porosity, specific surface area, tortuosity and constrictivity, have all been shown to be good predictors of transport and conductivity, see, for example, reference [15]. Specific surface area, for example, has a long history of being used in the context with the Kozeny–Carman equation. The fourth predictor set is chosen based on SHAP values computed for a collection of random forest models. Figure 14 clearly shows that through-intrusion porosimetry, tortuosity, effective porosity and sphere size compute in the pore space contribute most to permeability prediction for these fitted models, and so these four descriptors are chosen for the fourth set.

The fits of the first three predictor sets for the additive model are shown in Figure 15. First, we see that the model assumption of i.i.d. Gaussian noise $\mathbf{e}$ seems to hold well for all three fitted regression models (Panel A–C), as the data are homogeneously scattered around the diagonal line indicating the predicted value. The first regression model, shown in Panel A, includes only log-transformed effective porosity as predictor and fits fairly well. There is a relatively large portion of the variation in the data that remains unexplained however, as seen from the R^2-values in Table 3. Adding log-transformed specific surface area as a predictor (Panel B) improves the fit, and using all four log-transformed predictors (Panel C) gives further improvement. The difference is small between the R^2-values from fitting to the full dataset and R^2-values averaged over 10-fold cross-validation. This seems to be because the fitted smooth curves are insensitive to leaving out 10% of the data.

The fitted smooth function for each predictor when using the third predictor set (Figure 15C) are shown in Figure 15D–G. The fitted functions are almost linear, even though the model allows for non-

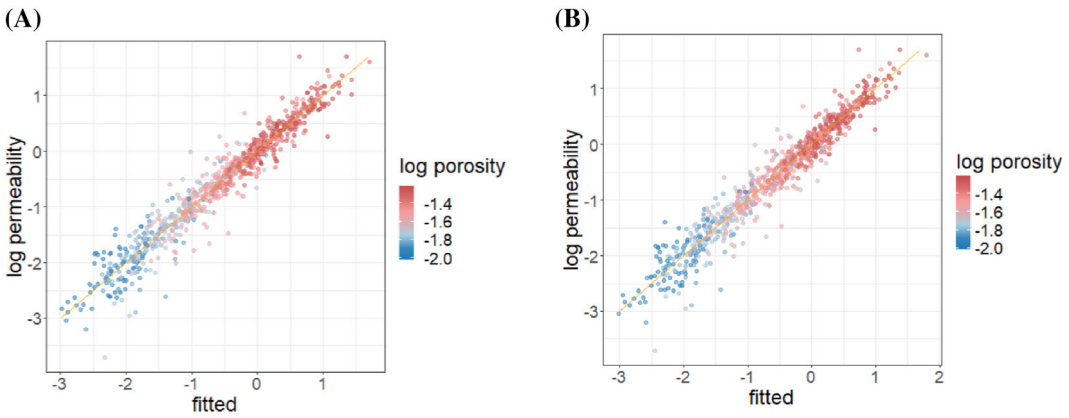

FIGURE 16 Relationship between the response variable log(permeability) and its fitted values, comparing the best two additive models with (A) effective porosity, specific surface area, tortuosity and constrictivity and (B) effective porosity, tortuosity, through-intrusion porosimetry and pore sphere size.

linear smooth functions. Thus it seems that a multivariate linear regression on log-scale would be a sufficient model. This is explored by fitting linear models for the four predictor sets. Results in Table 3 show almost identical R^2-values for the smooth and linear regression models. AIC values still indicate that the smooth regression models are an improvement over the linear model, at least for those models where the AIC values differ markedly.

The fourth predictor (Figure 16B) set performs almost exactly on par with the third set (Figure 16A) when evaluated with R^2, and slightly better when evaluated with AIC (Table 3). It is surprising that this model does not give a stronger improvement over the third predictor set as the SHAP values are higher overall in the fourth predictor set. A possible explanation could be the correlation between predictors, which results in individual descriptor contributions indicated by

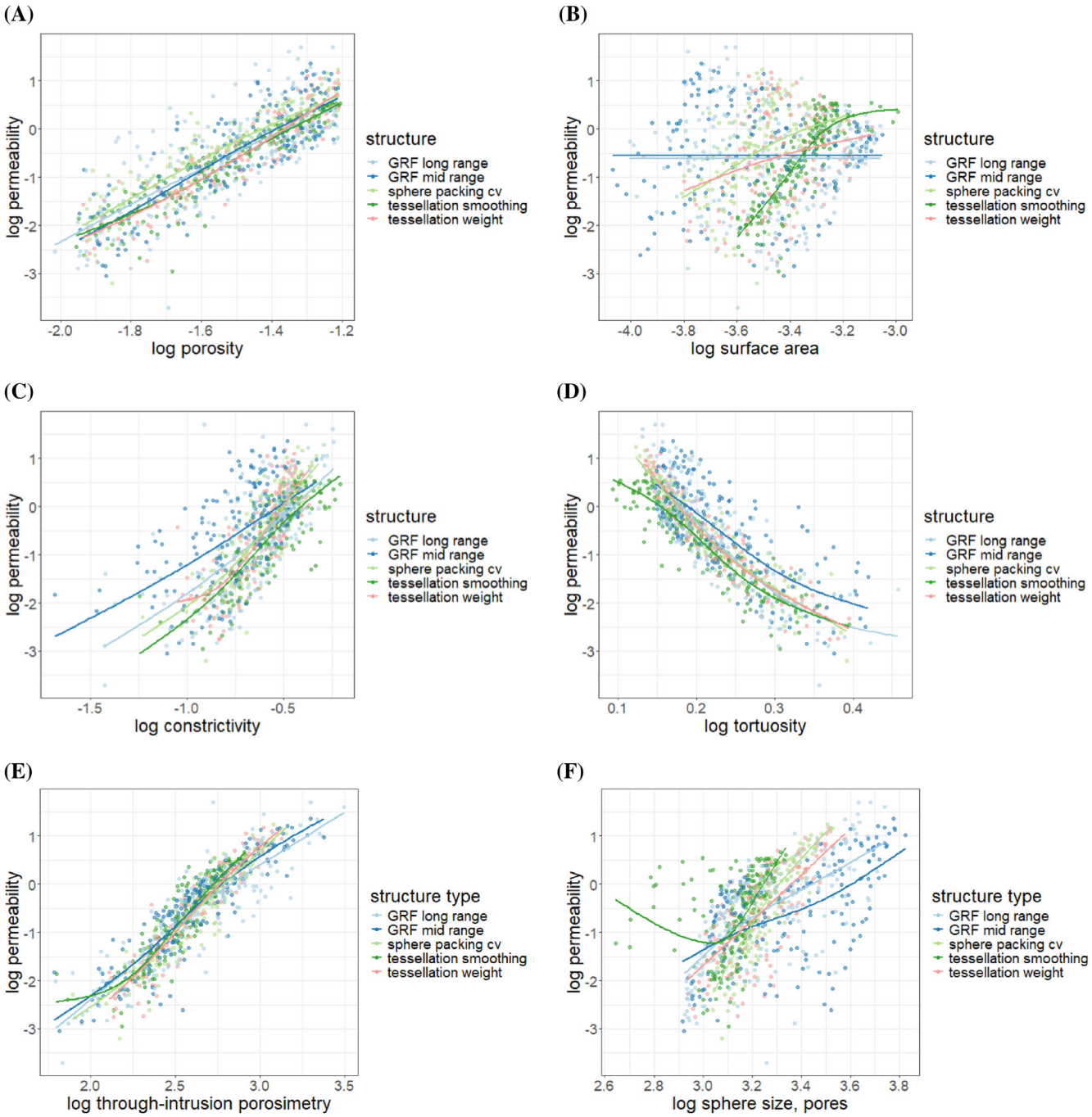

FIGURE 17 Relationship between log(permeability) and the microstructural descriptors (A) log(effective porosity), (B) log(specific surface area), (C) log(constrictivity), (D) log(tortuosity), (E) log(through-intrusion porosimetry) and (F) log(sphere size, pore). Each structure generated from the model is represented by one dot, and structure groups $D_{GRF,long}$ (GRF long range), $D_{GRF,mid}$ (GRF mid range), $D_{cv_{sph}}$ (sphere packing cv), D_{σ} (tessellation smoothing), D_w (tessellation weight) are represented with different colors. For each group, a smoothing curve representing the relationship between permeability and microstructural descriptor for that group is shown.

the high SHAP values not giving substantial improvements when used jointly. The fourth predictor set model can, as with the third predictor set, be interpreted in terms of porosity, path-length (tortuosity) and bottleneck effects. In the case of the fourth predictor set, the bottleneck effect is measured by the through-intrusion porosimetry whereas for the third predictor set it was measured by the constrictivity.

Figure 17 shows relationships between microstructural descriptors and permeability for descriptors included in the predictor sets, for each structure group $D_{cv_{sph}}$, D_{σ}, D_{w}, $D_{GRF,mid}$, $D_{GRF,long}$ separately. We see that the relationship between effective porosity ϵ and permeability γ (Figure 17A) is almost linear on the log scale, and similar for all five structure groups. The relationship between specific surface area and permeability (Figure 17B) on the log scale shows a higher degree of variation between the five groups than the relationship between the other descriptors and permeability (Figure 17A,C,D). The relationship between the constrictivity and permeability (Figure 17C) is similar for all five structure groups, showing an increasing permeability with increasing constrictivity. The same is true for the relationship between tortuosity and permeability (Figure 17D), but here the permeability decreases with increasing tortuosity. Through-intrusion porosimetry (Figure 17E) and sphere size computed in the pore space (Figure 17F) also show a high correlation with permeability, which is to be expected with their high SHAP values, although the relationship is more clear for through-intrusion porosimetry than for pore sphere size.

3 | CONCLUSIONS

We built a virtual material structure model for paperboard coatings and developed statistical methods for fitting this type of multi-component structure model to data. The model is a combination of a tessellation and a Gaussian random field, where the tessellation captures the large-scale and the random field the small-scale variation in the coatings. In the novel model fitting approach, the covariance of the GRF is first estimated using an appropriate subset of the data identified by the microstructural descriptor 3D sphere size and then, the remaining parameters are estimated using ABC.

The model fitting procedure, which was evaluated using simulated data, recovered the model parameters well. However, we could only identify the model parameters correctly when the 3D sphere size was included in the accept-reject discrepancy measure, see Section 2.4.2. From this we conclude that it is worth computing the 3D sphere size even though it is more computationally demanding to compute than the related contact distance.

The model fits well to the coating data. However, it is worth pointing out that we did not find any covariance function $C^{X_{GRF}}$, for which the model fits well both in terms of the 3D sphere size in the pore space and in terms of the sample covariance. As it is mainly the GRF that controls the 3D sphere size in the pore space, and the GRF is completely determined by its covariance, a non-stationary and/or non-Gaussian random field, which is determined not only by the covariance, might be able to overcome this problem. The way the model fitting is set up, parameter estimates are optimized for microstructural similarity with ABC. This algorithm could be tweaked by adding other discrepancy measures. The microstructural similarity would then be judged differently, and this could lead to a better fit for both the covariance and 3D sphere size. It would, however, likely not be sufficient to completely resolve the covariance/3D sphere size trade-off from what we have seen when using different ABC setups. Changing the random field model component is more likely to solve this issue.

We generated a wide range of microstructures from the model and explored how fluid permeability varies as we change the parameters of the virtual structure model. Using additive models with smooth non-linear functions relating microstructural descriptors to permeability, we showed that low porosity, strong bottleneck effects (low constrictivity/low through-intrusion porosimetry), long path-lengths (high geodesic tortuosity), and high specific surface area have a considerable limiting impact on the permeability through these structures. This shows how the microstructure of this specific model system relates to permeability through features directly interpretable regarding the microstructure. Efforts to optimize coating formulation can through these models evaluate how processing parameters impact microstructural features, such as how adjustments to calcium carbonate particle size distribution or binder content reduce bottlenecks effects or improves tortuosity, in that way optimizing directly those features which are controllable related to the microstructure and which in turn determine permeability.

AUTHOR CONTRIBUTIONS

Sandra Barman: Conceptualization; methodology; software; formal analysis; visualization; writing – original draft; writing – review and editing; validation; investigation. **Aila Särkkä:** Funding acquisition; conceptualization; methodology; validation; formal analysis; investigation; supervision; visualization; writing – review and editing.

ACKNOWLEDGMENTS

We acknowledge the financial support of VINNOVA (grant 2021-03665). The computations were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) and the Swedish National Infrastructure for Computing (SNIC) at the Chalmers Centre for Computational Science and Engineering (C3SE), partially funded by the Swedish Research Council (grant numbers 2022-06725 and 2018-05973).

CONFLICT OF INTEREST STATEMENT

M.R. is an employee of AstraZeneca and has stock ownership and/or stock options or interests in the company. The other authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

We are sharing all data we are able to share but we are restricted by our collaborative partner Tetra Pak to not share the FIB-SEM data of the coatings themselves. All other data related to model fitting and the simulation study, including code to perform the analysis, is shared

however. Code relating to model fitting and the simulation study is available and data relating to the simulation study is openly available at Zenodo repository.[36] Data from figures are when possible to share available in the Zenodo repository, along with code to recreate them.

ORCID

Sandra Barman https://orcid.org/0000-0002-4677-0794
Aila Särkkä https://orcid.org/0000-0003-4115-6530

REFERENCES

1. Fager C, Röding M, Olsson A, et al. Optimization of FIB-SEM tomography and reconstruction for soft, porous, and poorly conducting materials. *Microsc Microanal*. 2020;26:837-845.
2. Westhoff D, Manke I, Schmidt V. Generation of virtual lithium-ion battery electrode microstructures based on spatial stochastic modeling. *Comput Mater Sci*. 2018;151:53-64.
3. Barman S, Rootzén H, Bolin D. Prediction of diffusive transport through polymer films from characteristics of the pore geometry. *AIChE J*. 2019;65(1):446-457.
4. Townsend P, Larsson E, Karlson T, et al. Stochastic modelling of 3D fiber structures imaged with X-ray microtomography. *Comput Mater Sci*. 2021;194:110433.
5. Townsend P, Pingel TN, Lorén N, et al. Tessellation-based stochastic modelling of 3D coating structures imaged with FIB-SEM tomography. *Comput Mater Sci*. 2021;197:110611.
6. Hastie T, Tibshirani R. Generalized additive models. *Stat Sci*. 1986;1.3:297-310.
7. Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer; 2009.
8. Schindelin J, Arganda-Carreras I, Frise E, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676-682.
9. Schwartz J, Harris C, Pietryga J, et al. Real-time 3D analysis during electron tomography using tomviz. *Nat Commun*. 2022;13:4458.
10. Röding M, Fager C, Olsson A, von Corswant C, Olsson E, Lorén N. Three-dimensional reconstruction of porous polymer films from FIBSEM nanotomography data using random forests. *J Microsc*. 2021;281:76-86.
11. Ohser J, Mücklich F. *Statistical Analysis of Microstructures in Materials Science*. John Wiley & Sons; 2000.
12. Legland D, Kieu K, Devaux M-F. Computation of Minkowski measures on 2D and 3D binary images. *Image Anal Stereol*. 2007;26(2):83.
13. Chiu SN, Stoyan D, Kendall WS, Mecke J. *Stochastic Geometry and its Applications*. 3rd ed. John Wiley & sons; 2013.
14. Stenzel O, Pecho O, Holzer L, Neumann M, Schmidt V. Predicting effective conductivities based on geometric microstructure characteristics. *AIChE J*. 2016;62:1834-1843.
15. Prifling B, Röding M, Townsend P, Neumann M, Schmidt V. Large-scale statistical learning for mass transport prediction in porous materials using 90,000 artificially generated microstructures. *Front Mater*. 2021;8:786502. doi:10.3389/fmats.2021.786502
16. *MATLAB. Version R2023b*. The MathWorks Inc.; 2023.
17. Barman S, Gebäck T, Rootzén H, Bolin D, Olsson E, Särkkä A. *A Program Package for Visualization and Characterization of 3D Geometries*. MIST; 2019. https://mist.math.chalmers.se/.
18. Redenbach C, Zuyev S. Random Laguerre tessellations. *Adv Appl Probab*. 2008;40(9):630-650.
19. Williams CK, Rasmussen CE. *Gaussian Processes for Machine Learning*. MIT Press; 2006.
20. Adler RJ, Taylor JE. *Random Fields and Geometry*. Springer; 2009.
21. Lindgren F, Rue H, Lindström J. An explicit link between Gaussian random fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *J R Stat Soc Series B Stat Methodol*. 2011;73(4):423-498.
22. Rue H, Held L. *Gaussian Markov Random Fields: Theory and Applications*. CRC Press; 2005.
23. Schmidt V. *Stochastic Geometry, Spatial Statistics and Random Fields*. MIT Springer; 2014.
24. Cramér H, Leadbetter MR. *Stationary and Related Stochastic Processes*. John Wiley & Sons, Inc.; 1967.
25. Tavaré S, Balding DJ, Griffiths RC, Donnelly P. Inferring coalescence times from DNA sequence data. *Genetics*. 1997;145:505-518.
26. Pritchard JK, Seielstad MT, Perez-Lezaun A, Feldman MW. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Mol Biol Evol*. 1999;16:1791-1798.
27. Beaumont MA, Zhang W, Balding DJ. Approximate Bayesian computation in population genetics. *Genetics*. 2002;162:2025-2035.
28. Marjoram P, Molitor J, Plagnol V, Tavaré S. Markov chain Monte Carlo without likelihoods. *Proc Natl Acad Sci*. 2003;100:15324-15328.
29. Sisson SA, Fan Y, Tanaka MM. Sequential Monte Carlo without likelihoods. *Proc Natl Acad Sci*. 2007;104:1760-1765.
30. Sisson SA, Fan Y, Tanaka MM. Correction to "sequential Monte Carlo without likelihoods". *Proc Natl Acad Sci U S A*. 2009;106:16889.
31. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MPH. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *J R Soc Interface*. 2009;6:187-202.
32. Beaumont MA, Cornuet JM, Marin JM, Robert CP. Adaptive approximate Bayesian computation. *Biometrika*. 2009;96:983-990.
33. Gebäck T, Heintz A. A lattice Boltzmann method for the advection-diffusion equation with Neumann boundary conditions. *Comput Phys Commun*. 2014;15:487-505.
34. Gebäck T, Heintz A. A pore scale model for osmotic flow: homogenization and lattice Boltzmann simulations. *Transport Porous Media*. 2019;126(1):161-176.
35. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Statis Soc B*. 2011;73(1):3-36.
36. Barman S. Code, Computational Study of Permeability in Cardboard Coating Layers. Version v1. 2025. doi:10.5281/zenodo.14975592

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Barman S, Pingel TN, Lorén N, et al. Computational study of permeability in cardboard coating layers. *AIChE J*. 2026;72(7):e70388. doi:10.1002/aic.70388