



# Predicting Air Flow in Calendered Paper Sheets from $\mu$ -CT Data: Combining Physics with Morphology

Phillip Gräfensteiner<sup>1</sup> · Andoni Rodriguez<sup>2</sup> · Peter Leitl<sup>2,3</sup> · Ekaterina Baikova<sup>3,4</sup> · Maximilian Fuchs<sup>3,4</sup> · Eduardo Machado Charry<sup>3,4</sup> · Ulrich Hirn<sup>4,5</sup> · André Hilger<sup>6</sup> · Ingo Manke<sup>6</sup> · Robert Schennach<sup>3,4</sup> · Matthias Neumann<sup>7</sup> · Volker Schmidt<sup>1</sup> · Karin Zojer<sup>3,4</sup>

Received: 11 June 2025 / Accepted: 21 October 2025

© The Author(s) 2025

## Abstract

Predicting the macroscopic properties of thin fiber-based porous materials from their microscopic morphology remains challenging because of the structural heterogeneity of these materials. In this study, computational fluid dynamics simulations were performed to compute volume air flow based on tomographic image data of uncompressed and compressed paper sheets. To reduce computational demands, a pore network model was employed, allowing volume air flow to be approximated with less computational effort. To improve prediction accuracy, geometric descriptors of the pore space, such as porosity, surface area, median pore radius, and geodesic tortuosity, were combined with predictions of the pore network model. This integrated approach significantly improves the predictive power of the pore network model and indicates which aspects of the pore space morphology are not accurately represented within the pore network model. In particular, we illustrate that a high correlation among descriptors does not necessarily imply redundancy in a combined prediction.

---

✉ Phillip Gräfensteiner  
phillip.graefensteiner@uni-ulm.de

<sup>1</sup> Institute of Stochastics, Ulm University, Helmholtzstraße 18, 89069 Ulm, Germany

<sup>2</sup> bionic surface technologies GmbH, Liebenauer Hauptstraße 2–6, 8041 Graz, Austria

<sup>3</sup> Institute of Solid State Physics, NAWI Graz, Graz University of Technology, Petersgasse 16/II, 8010 Graz, Austria

<sup>4</sup> Christian Doppler Laboratory for Mass Transport through Paper, Graz University of Technology, Petersgasse 16/II, 8010 Graz, Austria

<sup>5</sup> Institute of Bioproducts and Paper Technology, Graz University of Technology, Inffeldgasse 23, 8010 Graz, Austria

<sup>6</sup> Institute of Applied Materials, Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin, Germany

<sup>7</sup> Institute of Statistics, Graz University of Technology, Kopernikusgasse 24/III, 8010 Graz, Austria

## Articlehighlights

- Air flow through paper sheets has been determined by computational fluid dynamics and pore network model simulations
- Geometric descriptors of pore space are used in regression models to improve the prediction of pore network model simulations
- Structure–property relationships indicate descriptors complementary to pore network modeling for different paper grades

**Keywords** Porous medium · Air permeance · Computational fluid dynamics · Pore network model · Statistical image analysis

## 1 Introduction

Predicting the macroscopic flow of gasses through a heterogeneous porous material requires knowledge of the size, shape, and connectivity of the pores involved, regardless of whether the flow is diffusive, laminar, or turbulent. Although methods for transport simulations exist that use 3D image data to take the geometry of the pore space into account (Torquato 2022), it remains difficult to establish relationships between morphology and transport for flow through thin, sheet-like structures. Prominent examples of such structures are fibrous membranes encountered in paper, gas diffusion layers, or filters. In such sheets, the pore structure often exhibits strong local heterogeneities. To understand the flow through the total sheet area, it is necessary to capture the local flow variations across many different realizations of the pore space. This requires an enormous amount of simulations, which, in turn, requires transport simulation methods that are accurate and, simultaneously, can be performed with reasonable computational effort.

The choice of the simulation method depends on whether diffusive, laminar, or turbulent transport is considered. In the case of laminar flow, there are several possible methods that, by and large, have in common that they represent different routes to predict the flow originating from the Navier–Stokes equation, reducing the transport problem to the simpler Stokes equation (Leal 2007). Computational fluid dynamics (CFD) and lattice Boltzmann simulations are suitable and well-established methods to determine local flows from the actual pore space Chung (2002; Succi et al. 1991). In CFD simulations, the pore phase, as supplied by tomographic image data, is represented by a surface mesh on which the Stokes equation is solved. This simulation method allows for readily considering specific situations such as stationary or incompressible flows. Lattice Boltzmann simulations track the motion by monitoring the velocity components along a given set of directions in a cubic grid. With increasing complexity of the pore space morphology, more velocity components and smaller time steps are required to accurately solve the Boltzmann equation that is equivalent to the stationary Navier–Stokes equation. Each of these methods is computationally demanding (Yang et al. 2016). On the other hand, ready-trained deep learning methods do not offer a route to evade the computational costs yet, as they exclusively act on the material classes they were trained for and cannot be universally applied to all types of materials, at least not without supplying costly simulations for retraining (Morgan and Jacobs 2020). However, the modeling of pore networks is a promising and computationally cheaper alternative (Blunt et al. 2013). In partitioning the pore

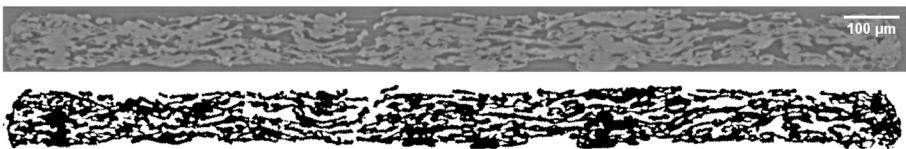
space into distinct components, pore network modeling (PNM) preserves the transport-relevant local connections and bottlenecks between the pores. Transport is considered locally by solving the transport equation through neighboring pores assuming a simplified geometry of involved pores. In order to estimate the total flow through the entire pore space, the local transport processes are consistently connected with each other by enforcing mass conservation.

It is highly desirable to interpret or even predict the fluxes using geometric descriptors associated with the pore space, as these can be reliably determined from tomographic image data without the need of performing physical transport simulations. Using methods from spatial statistics and mathematical morphology, it is possible to quantitatively characterize the complex morphology of the pore space (Chiu et al. 2013; Jeulin 2021; Ohser and Schladitz 2009). In particular, to quantify local heterogeneities in paper-based materials, methods have been developed to quantify the variance and correlation of multiple local geometric descriptors computed from tomographic image data (Neumann et al. 2024). This methodology has recently been applied to quantify local heterogeneities in polymer-based batteries (Neumann et al. 2022; Ademmer et al. 2023; Dodell et al. 2025).

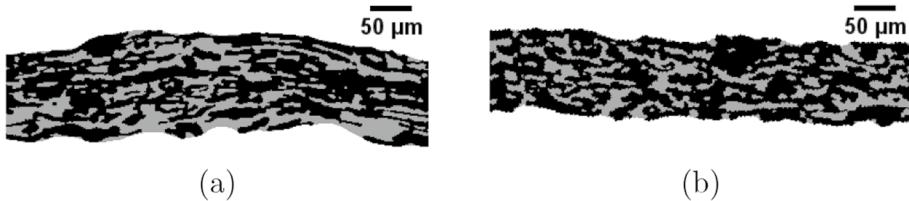
Predicting effective properties, such as permeability, based on such sets of descriptors is fundamental in many studies. The selected geometric descriptors of the pore space are usually inspired by the properties used to interpret the permeability of the material, for laminar flow in the framework of Darcy's law (Whitaker 1986). Prominent transport-relevant descriptors are the volume fraction of the pore space (porosity), the local thickness of the material, the internal surface area per unit volume, and the mean geodesic tortuosity to quantify the winding of transportation paths. However, determining how many and which descriptors are truly relevant for a given porous material is not straightforward. Particularly problematic is the commonly encountered situation that these descriptors depend on each other, i.e., varying the value of one descriptor immediately affects the values of all other descriptors. The specific consequences of correlations between descriptors depend on the material, since the actual degree of correlation between descriptors can vary strongly from material to material (Neumann et al. 2024).

In the present study, we investigate the problem of modeling laminar flow using the example of two paper samples with different morphologies. These paper sheets originate from the same paper grade whose structure is considered before and after compression in thickness direction. For both samples, the air fluxes have been determined experimentally and the microstructures have been acquired by means of  $\mu$ -CT measurements (Neumann et al. 2021). Local variations in the microstructures are quantified via transport-relevant descriptors of pore space that were computed based on the tomographic image data. Furthermore, the correlation structure between transport-relevant descriptors has already been quantified in previous works by means of R-vine copulas (Neumann et al. 2024, 2021).

For each paper sample, we relate the fluxes obtained from CFD simulations to the fluxes determined by PNM simulations and various geometric descriptors, using several power-law models. It must be emphasized that such fits do not suggest physically motivated expressions, i.e., the fitted exponents of the geometric descriptors cannot guarantee trends that reveal the underlying transport physics. However, by means of these models, we can study microstructure–property relationships between geometric descriptors of pore space and the volumetric flux as determined by CFD simulations. Furthermore, we can evaluate the quality of the PNM simulations and monitor which descriptors improve the prediction of effective properties and thus hint toward details in pore space morphology that are not captured in PNM simulations. In order to improve the agreement between the results of CFD and PNM simulations, we use combinations of various geometric descriptors of pore space, such as the mean geodesic tortuosity, the surface area per unit volume, or scalar quantities derived from the continuous pore size distribution.



**Fig. 1** Slices of raw gray scale (top) and binarized (bottom)  $\mu$ -CT data for compressed paper sheets



**Fig. 2** Slices of  $\mu$ -CT data for uncompressed (a) and compressed (b) paper sheets (from cutouts  $500\text{ }\mu\text{m} \times 500\text{ }\mu\text{m} \times \text{height}$ ). The solid and pore phase are depicted in black and gray, respectively

## 2 Materials, Tomographic Imaging, and Transport Regime

First, in Sect. 2.1, we give a short overview of the materials investigated in this study and the tomographic image data measured for these materials. In Sect. 2.2, we introduce the transport regime considered in the present study that will be solved by means of CFD and PNM simulations, as will be explained later in Sect. 3.

### 2.1 Materials and Their Tomographic Imaging

Our study is based on two data sets of tomographic image data for two different paper sheets. One sample type represents a paper sheet compressed in thickness direction, and one represents a sheet of the paper grade before compression (Neumann et al. 2021). The latter, uncompressed sample is commercial, unbleached paper with a specific basis weight of  $100\text{ g/m}^2$ . Compressed samples are obtained via hard-nip, steel–steel calendering of the uncompressed paper sheets with a line load of  $90\text{ Nm}^{-1}$ . The 3D microstructures of both samples are resolved by  $\mu$ -CT image data with a voxel size of  $1.3\text{ }\mu\text{m}$  as described in (Neumann et al. 2021). The binarization of the uncompressed paper sheets was performed based on absorption contrast using indicator (kriging Oh and Lindquist 1999) as described in (Machado Charry et al. 2018), whereas the binarization of the compressed paper sheets was obtained using a random forest classifier within the Fiji Weka segmentation plugin (Schindelin et al. 2012; Arganda-Carreras et al. 2017) as described in (Neumann et al. 2021), see Fig. 1.

For determining the boundary of the paper sheets, which particularly allows for computing the thickness, a rolling ball approach (Sternberg 1983) is used as already performed in (Neumann et al. 2022; Machado Charry et al. 2018). Exemplary slices of the segmented tomographic image data are shown in Fig. 2.

Since both types of (CFD and PNM) simulations performed in this study require spatially resolved 3D information of the pore phase, our common starting point is to predefine cutouts of the 3D stack of tomographic image data, where each cutout has a cross-sectional area of

$$A_{\text{seg}} = 500 \times 500 \mu\text{m}^2 \quad (1)$$

and contains all voxels in thickness direction. For each of the uncompressed and compressed paper sheets, we selected twelve cutouts to cover a large range of variations in the pore space morphology. In both cases, the cutouts were chosen such that eight of them have a local porosity close to the mean (global) porosity obtained for the entire microstructure data of the sample. Two cutouts represent realizations of particularly dense sheets (smaller porosity) and two further particularly open sheets (higher porosity). To facilitate a later interpretation of the simulation results, all cutouts have been chosen such that their mean thicknesses are approximately equal. Note that one of these cutouts from the compressed sample has not been used in the following analysis, because the CFD-determined porosity of this cutout was implausibly large and possibly indicates an artifact in the mesh generated from the microstructure, see Sect. 3.1 below.

## 2.2 Transport Regime

We turn to air flow in thickness direction through porous paper sheets under excess pressure. In this transport regime, we expect that chemical interactions of the molecules in air with the solid matrix are negligible. To get laminar flow rather than turbulent flow, the pressure difference between both sides of the sheet must be small enough. The pressure difference required for the standardized, experimental characterization of volume flows through paper sheets (Gurley test (ISO 5636-5:2013 2013)) is small enough to warrant laminar flow. Hence, we will use the pressure difference from the Gurley test in our flow simulations. This gives us the opportunity to check whether the simulations are in line with the experimentally obtained air fluxes (Leitl et al. 2023). Considering air transport rather than liquid further eases the comparison, because typical liquid-relevant complications such as capillary uptake, wetting in angular pores (Valvatne and Blunt 2004; Zhao et al. 2022) and associated slip-stick motion (Aslannejad et al. 2017; Fischer et al. 2021) do not appear.

### 2.2.1 Navier–Stokes Equation

We briefly show how the formulation of the Navier–Stokes equation of laminar transport of gases such as air can be simplified for the use of CFD transport simulations. Recall that the Navier–Stokes equation determines the time-dependent velocity field  $u : \Omega \times [0, \infty) \rightarrow \mathbb{R}^3$  on a given domain  $\Omega \subset \mathbb{R}^3$  under the influence of a given pressure field  $p : \Omega \times [0, \infty) \rightarrow [0, \infty)$  and negligible gravitational forces for each time  $t \geq 0$  and position  $x = (x_1, x_2, x_3) \in \Omega$ , where the domain  $\Omega$  contains the (connected) space of pores. It is common to add to  $\Omega$  additional void volumes above and below the sample, i.e., an inlet and outlet region, such that boundary conditions can be assumed on planar surfaces. As the considered pressure difference in our problem is high enough to assume incompressible flow (Bernabé 2018), the Navier–Stokes equation reads

$$\rho \left( \frac{\partial u}{\partial t} + (u \cdot \nabla) u \right) = -\nabla p + \eta \Delta u \quad \text{in } \Omega \times (0, \infty), \quad (2)$$

where  $\eta > 0$  is the dynamic viscosity and  $\rho > 0$  is the mass density (Sohr 2012). At the boundary between the pore and the solid phase, no-slip boundary conditions ensure that the velocity of the gas is equal to that of the rigid solid phase at the wall, which is 0. This also ensures that the flow through the pores cannot penetrate the solid matrix. As flows through porous media usually have very small Reynolds numbers, the so-called inertial contribution on the left-hand side of Eq. (2) is neglected. Thus, we assume that

$$\left( \frac{\partial}{\partial t} + (u \cdot \nabla) \right) u = 0 \quad \text{in } \Omega \times (0, \infty), \quad (3)$$

such that the steady-state Stokes equation

$$-\nabla p + \eta \Delta u = 0 \quad (4)$$

remains. Then, CFD simulations are performed to determine a solution of Eq. (4) together with the continuity equation

$$\nabla \cdot u = 0 \quad (5)$$

for a given pressure difference at the boundaries to obtain the velocity field  $u$  and the local pressure, see Sect. 3.1 below. The volume flow rate  $\partial V / \partial t$  through a plane  $A \subset \mathbb{R}^3$  that intersects the pore space  $\Omega$  is then given by

$$\frac{\partial V}{\partial t} = \int_{A \cap \Omega} \langle n_A, u \rangle \, d\mathcal{H}_2, \quad (6)$$

where  $n_A$  is the surface normal vector of  $A$  pointing in the direction of the flow,  $\langle \cdot, \cdot \rangle$  denotes the scalar product between two vectors, and  $\mathcal{H}_2$  is the two-dimensional Hausdorff measure. Note that we consider the steady state, in which the flow rate  $\partial V / \partial t$  does not depend on the time  $t$ . In the case of the total flow rate through a paper sheet in thickness direction, the volume flow rate is typically determined at a plane that is oriented normally to the pressure difference and is placed in the outlet region.

## 2.2.2 Hagen–Poiseuille Equation

Within the pore network model, the pore space is represented by a graph, in which connections between neighboring pores are geometrically modeled by a cylinder. The volume flow rate of every such connection is determined individually by assuming steady-state, laminar flow through a long, narrow pipe (Xu et al. 2022; Gombosi 1994; Zhao et al. 2020). In this setting, the flow rate can be modeled using the Hagen–Poiseuille equation (Hutten 2015). This yields that

$$\frac{\partial V}{\partial t} = \frac{\pi R^4}{8\eta L} \Delta p, \quad (7)$$

where  $\Delta p$  denotes the pressure difference between the two ends of the pipe,  $\eta > 0$  is the dynamic viscosity,  $L > 0$  is the length of the pipe, and  $R > 0$  is the pipe radius. Note that the volume flow rate  $\partial V / \partial t$  of a given segment is directly proportional to the difference in pressure between the beginning and the end of the segment. Finally, the flow through the entire network must obey a mass balance equation at every vertex. For more details on how

the graph is constructed in the pore network model and how the linear factors in Eq. (7) are determined, see Sect. 3.2.

### 3 Methods to Simulate Transport

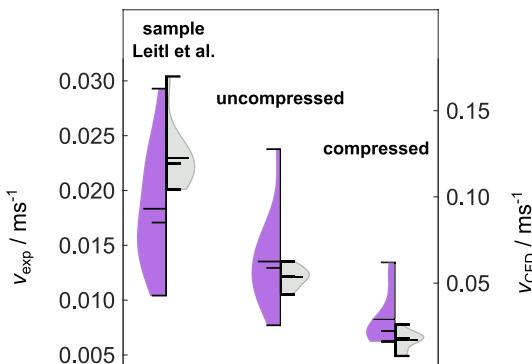
We now provide more detailed information on the two approaches used to simulate volume air flow through paper sheets. In Sect. 3.1, we describe the CFD simulation procedure and compare the resulting values to experimental data. In Sect. 3.2, the pore network model is described in multiple steps. First, Sect. 3.2.1 describes the construction of the graph representing the pore space morphology. Then, in Sect. 3.2.2, we show how the graph is used for volume air flow simulations. Sect. 3.2.3 discusses the impact of the chosen conduit shape used to approximate the volume between two neighboring pores within the pore network model.

#### 3.1 CFD Simulations

The ANSYS code was used to numerically solve the Stokes equation (see Eq. (4)) in the pore space of the paper sheets as described in Sect. 2.2, where the simulations were prepared as follows (Leitl et al. 2023). The microstructure of each cutout was provided as a stack of 2D binary images containing voxels corresponding either to the pore or to the solid phase. This volume data was triangulated to obtain the surface mesh of the pore space, since this surface mesh encloses the pore volume in which the Stokes equation will be solved. To arrive at a good compromise between high resolution and uniform coverage of surface triangles, we stretched the volume data in thickness direction by a factor of five, loaded this stretched volume data as images in (Fiji Schindelin et al. 2012) to perform the triangulation, and stretched the resulting surface mesh by a factor of 0.2 in thickness direction. This is done in order to compensate for the difference in lengths between thickness and lateral direction, and arrive at a surface triangulation whose typical side lengths are more proportional to the difference between thickness and lateral size. As ANSYS relies on a finite volume approach, the pore volume enclosed by the surface mesh was discretized using Numeca Hexpress Hybride. More details can be found in (Leitl 2020). The surface meshing closes all pores that are not connected to the exterior, i.e., either to an inlet or outlet. Air permeance was obtained for boundary conditions that correspond to the standardized test to assess the porosity of paper (ISO 5636-5:2013 2013). The pressure at the top surface was  $p_{\text{inlet}} = 102.545 \text{ kPa}$ , and  $p_{\text{outlet}} = 101.325 \text{ kPa}$  at the bottom surface. At the remaining side walls, symmetry boundary conditions are used. Also the external conditions were chosen to match the standardized conditions. The temperature was  $T = 298 \text{ K}$ , so that the density of air was  $\rho_{\text{air}} = 1.184 \text{ g m}^{-3}$ , and the dynamic viscosity was  $\eta = 1.838 \cdot 10^{-5} \text{ Pas}$ . For each cutout, the CFD simulations yield the spatially resolved air pressure and air velocity, and at the outlet side the volume flow rate  $\partial V / \partial t$  (via Eq. (6)) and the volume flux per unit area  $v_{\text{CFD}}$ , which is given by

$$v_{\text{CFD}} = \frac{1}{A_{\text{seg}}} \frac{\partial V}{\partial t}, \quad (8)$$

where  $A_{\text{seg}}$  is defined as in Eq. (1).



**Fig. 3** Violin plots to compare the fluxes obtained experimentally by the Gurley method (purple, left axis) and the fluxes obtained from CFD simulations (gray, right axis) for a paper sample from (Leitl et al. 2023), uncompressed, and compressed paper. To ease the comparison, plots of corresponding fluxes are superimposed at the common symmetry axis of the violin plot and a half of each violin plot is hidden. In each plot, the short horizontal bars represent the median and the quartiles, and the long horizontal bar the mean

Figure 3 compares the volume fluxes per unit area computed in this way with the fluxes  $v_{\text{exp}}$  obtained from experiments for uncompressed (center) and compressed paper (right). In the experiment, the time required to press a standard volume of 100mL air through a paper sheet of a standardized area of  $645.2 \text{ mm}^2$  is measured. (According to norm TAPPI T 460, this corresponds to a 1.0 square inch circular area.) For each paper grade, these times have to be measured at least five times at different positions and again at least five times with flow penetrating the paper in the opposite direction, i.e., with the paper sheet flipped. The measured times  $t_G$  (in seconds) are then converted to volume flow rates per unit area via  $v_{\text{exp}} = 0.155 \text{ m} / t_G$ . As the measured area of this experiment largely exceeds the area of the cutouts in the CFD simulations, the variations seen in the experiment stem from probing different positions. The fluxes  $v_{\text{exp}}$  and  $v_{\text{CFD}}$  differ by a scaling factor between four and five, which is consistent with a previous comparison for a paper made from the same pulp and a porosity comparable to the uncompressed sample (Leitl et al. 2023), whose simulated and experimentally obtained fluxes are shown for comparison in the left part of Fig. 3. However, the trends seen in the experiments are captured by the simulations for both samples, in terms of mean values and variations. Hence, in the context of the present study, we consider the CFD-calculated fluxes as the ground truth. They will serve as a reference for the pore network simulations.

### 3.2 PNM Simulations

The PNM approach seeks to predict transport in a geometrically simplified, graph-like representation of the pore space. The pore network representation of a pore space contains the positions and sizes of distinguishable pore regions, and the positions and sizes of the cross-sectional contact areas between neighboring regions.

In the network, each vertex represents the center of a pore region. Vertices corresponding to centers of connected pore regions receive an edge in the graph. A pore space analysis step, described in detail in Sect. S.3 of the Supplementary Information, determines and labels the pore regions and augments the vertices in the graph with information such as the

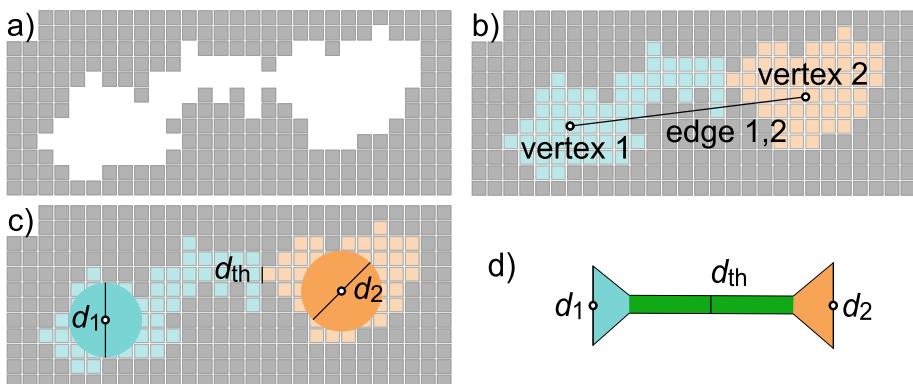
position (center of gravity) and volume of the original pore, and the diameter of the largest sphere that can be inscribed into the pore. Each edge is marked with information related to the “bottleneck” between two connected pores, i.e., the position and extension of the contact area between neighboring pore regions.

The transport between two vertices that are connected by an edge is modeled by solving a transport problem on an approximating conduit shape that is constructed based on the geometric information of each vertex and edge mentioned above. Correspondingly, the transport simulation is approximated in two steps: (i) determine a pore network representation of the pore space and (ii) perform simulation of Stokes flow on neighboring pores by use of approximating conduit shapes. These steps are explained in Sects. 3.2.1 and 3.2.2, respectively. Section 3.2.3 discusses the impact of the conduit shape on the fluxes predicted by PNM.

### 3.2.1 Constructing the Pore Network Graph

The pore network of each cutout was determined with the SNOW algorithm as implemented in the python package PoreSpy (Gostick et al. 2019). First, the SNOW algorithm divides the pore space of the segmented 3D image into non-overlapping regions, see Fig. 4a–c. Subsequently, an undirected geometric graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  is constructed, where each vertex  $\mu \in \mathcal{V} \subset \mathbb{R}^3$  corresponds to the centroid of a region in the segmented image, see Fig. 4b. Then, edges  $e = \{\mu, v\} \in \mathcal{E}$  are added between vertices  $\mu$  and  $v$  corresponding to neighboring regions, i.e., between those that share a common boundary.

These vertices and edges receive geometric information deduced from the tomographic image data of the pore regions. The volume of a pore region is the total volume of all voxels in a region. The associated pore radius is the maximum value of the Euclidean distance map (Soille 2003; Maurer et al. 2003) within each pore region, i.e., it corresponds to the radius of the largest possible sphere that is fully contained in the pore region. The pore surface area is given by the number of solid–pore interface voxels of the region multiplied



**Fig. 4** Pore space before (a) and after (b) partitioning in distinct pore regions, where each pore region is represented by a vertex, and edges mark adjacent, connected pore regions. Analysis of the pore regions provides the positions of the pores, the diameters  $d_1$  and  $d_2$  of inscribed spheres of maximum diameter, the positions in which the pore regions touch and the largest Euclidean distance  $d_{th}$  therein (c). Example for a straight pore-throat-pore conduit between two connected vertices in the pore network (d), where the conduit consists of the pores simplified as truncated cones, whose inlet and outlet diameters are given by  $d_1$  and  $d_2$ , respectively, and a central cylinder with diameter  $d_{th}$

by the area of a single voxel face. Each edge receives a diameter that corresponds to the largest Euclidean distance found in the cross-sectional area of two adjacent pore regions and a length, that corresponds to the Euclidean distance between the two connected pore centroids (Gostick 2017), see Fig. 4c. A more detailed description of how the pore network graph was determined is given in Sect. S.3 of the Supplementary Information.

While the pore network graph is constructed based on the morphology of the original pore space of the material, the geometry of the resulting graph differs markedly from the original microstructure due to its simplified representation. In fact, even defining a porosity of the pore network graph is difficult, as information the solid phase is completely discarded. A more detailed discussion on this can be found in Sect. S.4 of the Supplementary Information.

### 3.2.2 Flow Simulation on the Network

Based on the pore network graphs stated above, the OpenPNM (Gostick et al. 2016) package was used to determine the stationary air flow for the same excess pressure and external conditions as used for the CFD simulations (ISO 5636-5:2013 2013). The flux through the network has to obey the mass balance equation at every vertex  $\mu \in \mathcal{V}$  in the graph (Ioannidis and Chatzis 1993), i.e., for the net flow rate  $(\partial V / \partial t)_\mu$  through vertex  $\mu$  it holds that

$$\left( \frac{\partial V}{\partial t} \right)_\mu = \sum_{v \in N_\mu} \left( \frac{\partial V}{\partial t} \right)_{\mu v} = 0, \quad (9)$$

where  $N_\mu = \{v \in \mathcal{V} : \{\mu, v\} \in \mathcal{E}\}$  is the set of neighbors of vertex  $\mu$  and  $(\partial V / \partial t)_{\mu v}$  is the flow rate between adjacent vertices  $\mu$  and  $v$ , which is positive when the flow is directed toward the vertex  $\mu$ , and negative when the flow leaves the vertex  $\mu$ .

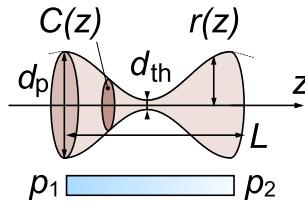
We assume here that the gas behaves like an incompressible liquid with a constant mass density. The disregard of compressibility has been shown to have little impact provided that the excess pressure driving the flow is high enough (Bernabé 2018). Hence, conserving mass is equivalent to conserving volume.

The flow between adjacent vertices  $\mu, v \in \mathcal{V}$  is driven by the difference in pressures  $p_\mu - p_v$  within the pore regions and adopts a form inspired by the Hagen–Poiseuille equation for laminar flow in pipes (Xu et al. 2022; Gombosi 1994; Zhao et al. 2020) given by

$$\left( \frac{\partial V}{\partial t} \right)_{\mu v} = g_{\mu v} (p_\mu - p_v), \quad (10)$$

where  $g_{\mu v} > 0$  is the local conductivity of the conduit between the adjacent vertices  $\mu$  and  $v$ .

Note that the local conductivity depends on the size and the geometric shape of the conduit. More precisely, the local conductivity (denoted by  $g$  in this paragraph) quantifies the ease with which a gas or fluid can pass through a conduit. For the hydraulic conductance considered here,  $g$  is fully determined by the shape of the conduit and the dynamic viscosity  $\eta$ . Fig. 5 illustrates a general, axial-symmetric conduit with a flow along the  $z$ -axis. We assume that the conduit is placed within the positive quadrant of the coordinate system, with one of its sides aligned with the  $x$ - $y$ -plane, so that its extension along the  $z$ -axis is from 0 to  $L$  for some length  $L > 0$ . For each  $z \in [0, L]$ , let  $C(z) \subset \mathbb{R}^3$  be the circular conduit cross-section perpendicular to the flow axis, and  $\mathcal{H}_2(C(z))$  the area of  $C(z)$ , given by means of the two-dimensional Hausdorff measure  $\mathcal{H}_2$ . Then, it holds that



**Fig. 5** Axial-symmetric conduit with varying radius  $r(z)$  and cross section  $C(z)$  along the  $z$ -axis. The inlet is at pressure  $p_1$  and the outlet at  $p_2$ , respectively. The diameter of the inlet and the outlet corresponds to the diameter  $d_p$  of the inlet and outlet pore, respectively. The smallest diameter is given by the extension of the bottleneck  $d_{th}$  between inlet and outlet pore

$$\frac{1}{g} = 16\pi^2\eta \int_0^L \frac{I_p^*(z)}{\mathcal{H}_2(C(z))^2} dz, \quad (11)$$

where  $I_p^*(z)$  is the specific polar moment of inertia, defined as integral over the cross section  $C(z)$  by

$$I_p^*(z) = \frac{1}{\mathcal{H}_2(C(z))} \int_{C(z)} \|x\|^2 d\mathcal{H}_2(x), \quad (12)$$

where  $\|x\|$  denotes the Euclidean norm of  $x = (x_1, x_2) \in \mathbb{R}^2$ . In Eq. (11), both  $\mathcal{H}_2(C(z))$  and  $I_p^*$  essentially depend on the geometry of the conduit. Thus, analytical descriptions of the shape of the cross sections along the flow axis often allow to derive analytical expressions for the conduit flow (Sochi 2013; Akbari et al. 2011). More details on this are given in Sect. S.5 of the Supplementary Information.

In this study, we assume that the conduit shape is the same throughout the network, where we choose the shape shown in Fig. 4d, i.e., a cone-cylinder-cone conduit consisting of three segments. The inlet segment starts from the pore center with opening diameter  $d_1$  and narrows in a conical fashion until the smallest diameter  $d_{th}$  is reached (pore). A second (straight) segment of cylinder shape with diameter  $d_{th}$  (throat) follows and connects to a conical-shaped outlet segment with smallest diameter  $d_{th}$  and final diameter  $d_2$ . Note that conduit shapes can be selected from a wide range of shapes as long as local conductivity through the conduit can be provided to satisfy Eq. (10), see (Zhao et al. 2020; Sochi 2013; Akbari et al. 2011; Miao et al. 2017) for examples. A more detailed discussion of the choice of conduit shape will be given in Sect. 3.2.3 below.

For a conduit between connected pores (represented by the vertices  $\mu, \nu \in \mathcal{V}$ ) consisting of multiple elements, such as the one shown in Fig. 4d, the local conductivity  $g_{\mu\nu}$  is obtained from the conductivities of the individual conduit elements (Rodriguez de Castro et al. 2023), i.e.,

$$\frac{1}{g_{\mu\nu}} = \frac{1}{g_{P,\mu}} + \frac{1}{g_{th,\mu\nu}} + \frac{1}{g_{P,\nu}}, \quad (13)$$

with  $g_{P,\mu}$  and  $g_{P,\nu}$  being the conductivities associated to the pore volumes (half-cones in Fig. 4d), and  $g_{th,\mu\nu}$  is the conductivity of the throat, where all three conductivities  $g_{P,\mu}$ ,  $g_{P,\nu}$  and  $g_{th,\mu\nu}$  are determined by Eq. (11).

With the local conductivities  $g_{\mu\nu}$  in hand, Eq. (9) defines a system of linear equations that is solved for the pressure  $p_\mu$  in each pore with the boundary conditions  $p_{\text{inlet}} = 1.22 \text{ kPa}$

at all inlet pores and  $p_{\text{outlet}} = 0$  at all outlet pores. Once the pressures  $p_\mu$  are determined, Eq. (9) provides the flow rates through each conduit. Then, the total flow rate through the network is the sum over all local volume flow rates entering the network at the inlet pores, i.e.,

$$\left(\frac{\partial V}{\partial t}\right)_{\text{inlet}} = \sum_{\mu \in \mathcal{V}_{\text{inlet}}} \left(\frac{\partial V}{\partial t}\right)_\mu = \sum_{\mu \in \mathcal{V}_{\text{inlet}}} \sum_{\nu \in N_\mu} g_{\mu\nu} (p_\nu - p_\mu), \quad (14)$$

where  $\mathcal{V}_{\text{inlet}} \subset \mathcal{V}$  is the set of all inlet vertices, which are determined as described in Sect. S.3 of the Supplementary Information. In analogy to Eq. (8), the volume flux per unit area  $v_{\text{PNM}}$  is given by

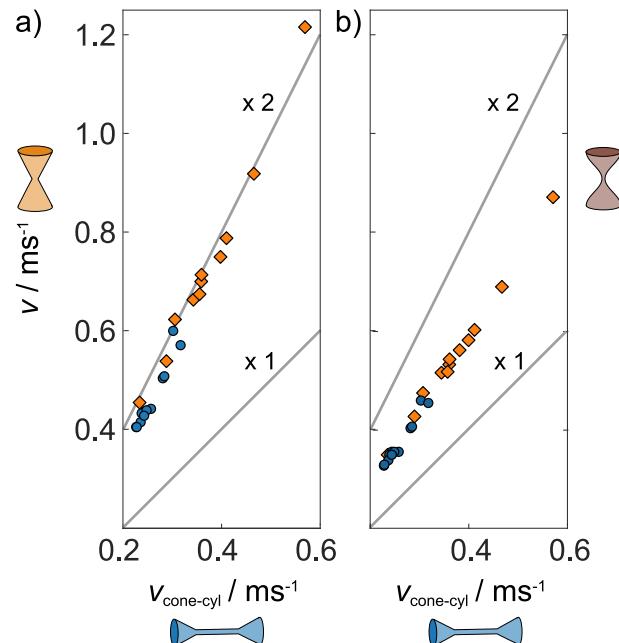
$$v_{\text{PNM}} = \frac{1}{A_{\text{seg}}} \left(\frac{\partial V}{\partial t}\right)_{\text{inlet}}. \quad (15)$$

### 3.2.3 Impact of Conduit Shape

Since the true morphology of the pore space region that connects two different pores is too complex, it must be approximated by using a simplified geometry in PNM to simulate the flow between neighboring pores. As there are many options for doing so, the values of local conductivities and, hence, of the overall flux  $v_{\text{PNM}}$ , depend on the choice of the conduit shape, see Fig. 6.

To assess the extent to which the predicted flows may vary, we consider the results of our simulations using the conduit shape shown in Fig. 4d as reference and compare the related fluxes to pore network simulations using two other conduit shapes. The associated

**Fig. 6** Impact of the conduit shapes on the fluxes predicted by PNM for the cutouts of uncompressed (orange diamonds) and compressed (blue circles) paper. The fluxes obtained for half pore-throat-half pore conduits as shown in Fig. 4d (horizontal axis) are compared to predictions using mirror- and axis-symmetric converging-diverging conduits with diameters  $d_{\text{max}} = d_i$  and  $d_{\text{min}} = d_{\text{th}}$  assuming a cone **a** and hyperbolic cosine **b** shape. To ease the comparison, two additional lines corresponding to a direct correspondence  $v = v_{\text{cone-cyl}}$  and a doubling in value  $v = 2v_{\text{cone-cyl}}$  are inserted



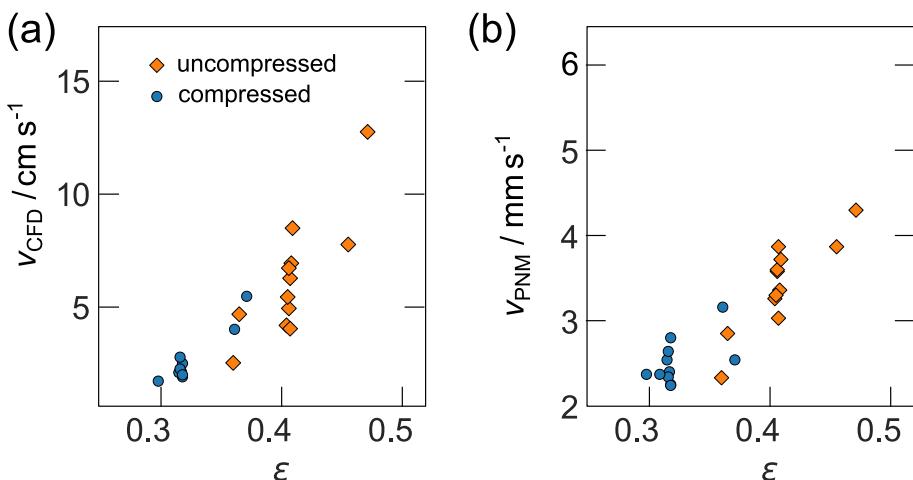
conduits assume a converging–diverging shape that are chosen to be mirror-symmetric, i.e., the diameters along the conduit fulfill  $d_{v,\max} = d_{\mu,\max} = d_{P,v}$  and  $d_{\min} = d_{\text{th},\mu v}$ . Using the notation  $v$  instead of  $v_{\text{PNM}}$ , Figure 6a shows the flux values of  $v$  obtained for conically shaped conduits for uncompressed (blue diamonds) and compressed (orange circles) paper, and compares them to the reference values (denoted by  $v_{\text{cone-cyl}}$ ) provided on the horizontal axis.

Although conically shaped conduits tend to show the largest deviations from cylindrical conduits (cf. Fig. 6a), the obtained values of  $v$  are practically proportional to the reference values of  $v_{\text{cone-cyl}}$  with a slope of two. The same qualitative behavior is obtained for hyperbolic cosine shaped conduits, see Fig. 6b. Also here, the values of  $v$  are proportional to the reference values of  $v_{\text{cone-cyl}}$ , but with a reduced slope of 1.5.

The flux through individual conduits of conical and hyperbolic cosine shapes is compared to that of cylindrical conduits in Figure S4 of the Supplementary Information.

### 3.3 Volume Flow Rates Obtained by CFD and PNM Simulations

In this section, we compare the volume flow rates per unit area obtained from CFD and PNM simulations, see Fig. 7. The values of the volume flow rates  $v_{\text{CFD}}$  and  $v_{\text{PNM}}$  differ by more than an order of magnitude. The actual difference in magnitude is determined, at least in part, by the choice of the conduit geometry, as explained above. Nevertheless, the flows predicted by PNM resemble the trends in the flows obtained by CFD. While the fluxes through the uncompressed sample tend to exceed the fluxes in the compressed sample, there is a common pattern regardless of which sample is considered: A higher porosity tends to give a larger flux. For comparable porosities, CFD and PNM simulations predict a marked spread in the flow values, see Fig. 7. This spread in flow rates must originate from the details of the pathways realized in each of the cutouts. Hence the question arises, whether other microstructure descriptors are capable of explaining this spread because they inherently consider these pathway details.



**Fig. 7** Volume flux per unit area obtained by CFD (a) and PNM (b) simulations versus the local porosity of the cutouts from the 3D image data for uncompressed (diamonds) and compressed (circles) paper

## 4 Geometric Descriptors of Pore Space Morphology

Based on the segmented  $\mu$ -CT image data, we can quantitatively analyze the pore space morphology within the measured paper sheets by means of various geometric microstructure descriptors. In particular, in a previous study (Neumann et al. 2024), we showed that the differences between uncompressed and compressed paper are heavily reflected in an altered interdependence between pairs of geometric descriptors of the pore space morphology. As we now aim at linking descriptors that characterize the pore space morphology to the volume flow determined by CFD simulations, as described in Sect. 3.1, we select an appropriate set of these geometric descriptors. Namely, we consider the porosity  $\varepsilon$ , the specific surface area  $S$  of pore space, the mean value  $\mu(\tau)$  and standard deviation  $\sigma(\tau)$  of geodesic tortuosity of paths through the pore space, and the median radius  $r_{\max}$  of the continuous pore size distribution. It is well known that each of these descriptors is relevant for transport in porous media (Holzer et al. 2013; Barman et al. 2019; Neumann et al. 2020).

### 4.1 Computation of Geometric Descriptors

In the following, we briefly explain each of the geometric descriptors considered in the present study and give information on how they can be computed from grid-based 3D image data.

**Porosity** The most fundamental and widely used geometric descriptor of pore space is the porosity  $\varepsilon \in [0, 1]$ . Formally,  $\varepsilon$  is defined as the volume fraction of the pore space, which can be determined by computing the number of voxels associated with pores divided by the total number of voxels that do not belong to the background. Thus,  $\varepsilon$  can be computed in a straightforward manner, simply by counting voxels in the image data.

**Specific surface area** Another fundamental descriptor is the surface area of pore space. In particular, we consider the specific surface area  $S$ , i.e., the surface area of pore space per unit volume. In order to compute the value of  $S$  from voxelized image data, an algorithm stated in (Schladitz et al. 2006) is used, which is based on local weighted  $2 \times 2 \times 2$  configurations. More precisely, the original binary image is convoluted with a  $2 \times 2 \times 2$  mask, resulting in a gray scale image where each possible gray scale value corresponds to a unique configuration of the  $2 \times 2 \times 2$  neighborhood of that voxel. These configurations have their own weightings, which are then summed up over the whole image in order to obtain an estimate for the surface area.

**Geodesic tortuosity** The general notion of tortuosity aims to quantify the length of transportation paths through a porous medium in relation to its thickness. Note that there are many different definitions of tortuosity (Holzer et al. 2023). However, in the present study, we focus on the concept of geodesic tortuosity. This requires the selection of starting and target planes of the image data, for which the geodesic tortuosity shall be computed. As we are interested in flow that traverses the paper sample vertically, we chose our starting and target planes as the pore space voxels that belong to the upper- and lower-most layers of voxels in  $y$ -direction, respectively. In addition, our goal is to quantify only the lengths of paths that use pores with a certain minimum local volume, so that we ensure that the considered paths contribute to volume flow in a significant way. We therefore first determine the pore space that can be filled by spheres of radius  $1.5 \mu\text{m}$ . This value was found to be suitable in a previous study (Neumann et al. 2021). Then, for every pore voxel of the starting plane, a shortest path to the target plane is computed twice by the use of

Dijkstra's algorithm (Dijkstra 1959), where once the path is only allowed to traverse the pore space, while the second time it is also allowed to pass through voxels that belong to the solid phase. The geodesic tortuosity  $\tau \geq 1$  is then defined as the ratio of the lengths of these two shortest paths for every pore voxel of the starting plane. This yields a distribution of values, of which we consider the sample mean  $\mu(\tau)$  and sample standard deviation  $\sigma(\tau)$ .

**Continuous pore size distribution** In order to capture the distribution of pore widths across the sample, we consider the continuous pore size distribution CPSD :  $[0, \infty) \rightarrow [0, 1]$ , which is defined using morphological opening (Soille 2003; Serra 1982). More precisely, for each  $r \in [0, \infty)$ , the value of CPSD( $r$ ) is given by the volume fraction of that part of the pore space that can be covered by spheres of radius  $r$  which are completely contained in the pore space. To compute CPSD( $r$ ) from voxelized image data, we consider the subset  $P \subset W$  of voxels associated with pores within our sampling window  $W \subset \mathbb{Z}^3$ . Then, we put

$$\text{CPSD}(r) = \frac{|(P \ominus B_{\mathbb{Z}^3}(o, r)) \oplus B_{\mathbb{Z}^3}(o, r)|}{|P|}, \quad (16)$$

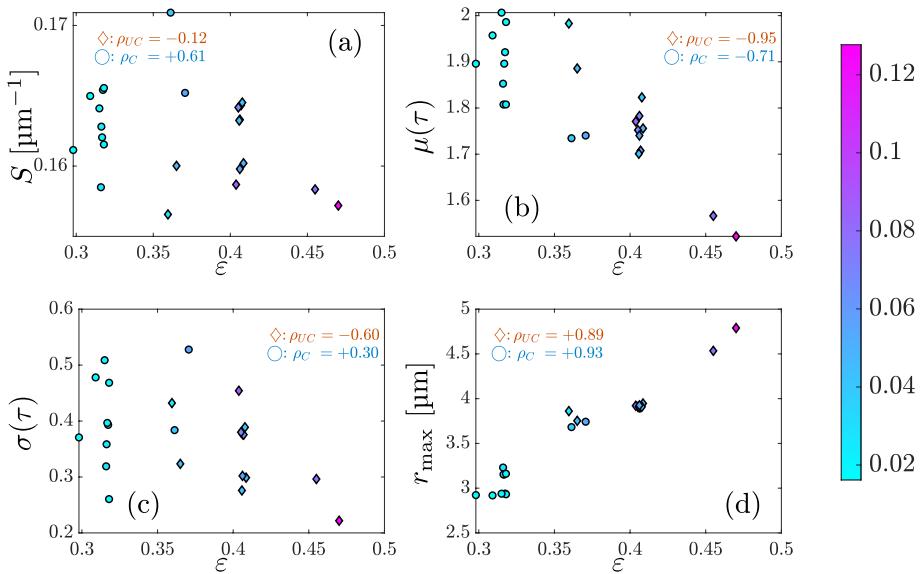
for any  $r \geq 0$ , where  $\ominus$  and  $\oplus$  denote the morphological operations of erosion and dilation (Soille 2003; Serra 1982), respectively,  $B_{\mathbb{Z}^3}(o, r) = B(o, r) \cap \mathbb{Z}^3$  is the discretized open ball of radius  $r$  centered at the origin, and  $|B|$  denotes the number of elements of any set  $B \subset \mathbb{Z}^3$ . Note that the numerator on the right-hand side of Eq. (16) can be efficiently computed by means of the Euclidean distance transform (Soille 2003; Maurer et al. 2003). In particular, we are interested in the median  $r_{\max}$  of CPSD, which is given by  $r_{\max} = \max\{r \geq 0 : \text{CPSD}(r) \geq 1/2\}$ . This quantity is used in the definition of constrictivity, which is a useful descriptor in quantifying the strength of bottleneck effects (Holzer et al. 2013).

## 4.2 Correlations Between Pairs of Geometric Descriptors

Figure 8 shows scatter plots of the results which we obtained for the geometric descriptors stated in Sect. 4.1, for the cutouts of both uncompressed and compressed paper sheets. It also shows the Pearson correlation coefficients for each pair of displayed descriptors, where we can see that some descriptor pairs show a rather strong correlation. Namely, the pairs  $\varepsilon$  and  $\mu(\tau)$  (Fig. 8b) as well as  $\varepsilon$  and  $r_{\max}$  (Fig. 8d) are strongly correlated in both samples.

In line with this, the associated fluxes, imposed by color in Fig. 8, tend to be higher the higher the porosity  $\varepsilon$  and the median radius  $r_{\max}$ , and the lower the mean geodesic tortuosity  $\mu(\tau)$ , which can be seen best in Fig. 8b and d. In contrast to this,  $\varepsilon$  and  $S$  only show a significant correlation for the cutouts of the compressed sample, see Fig. 8a, as also found in our previous work (Neumann et al. 2024).

Figure 8 also shows that at least half of the data points per sample share a practically identical local porosity  $\varepsilon$ . Although our pick of cutouts with similar local porosity may adversely affect the estimation of the Pearson correlation coefficient for the interrelations representing the whole sample, it will boost the importance of the other geometric descriptors in explaining variations in the flow. Note that the fact that the geometric descriptors considered in the present study are correlated affects the variance and interpretability of the estimated coefficients in the regression models that we will present in Sect. 5.1 below.



**Fig. 8** Scatter plots visualizing the interdependence of geometric descriptors for the cutouts of uncompressed (diamonds) and compressed (circles) paper sheets. The color coding indicates the values of the volume flow obtained by CFD simulations. Additionally, the corresponding values of the Pearson correlation coefficient are displayed on each figure

However, the presence of correlation among explanatory variables does not in general influence the ability to obtain a good fit (Kutner et al. 2004), see also Sect. 6.1.

### 4.3 Influence of the Characteristic Bottleneck

While the descriptors introduced in Sect. 4.1 are both easy to compute and also intuitive to interpret and to link to effective transport properties, the list is clearly not exhaustive and there exist many further interesting geometric descriptors that are not considered in our regression models, but can still be linked to fluid flow. In particular, the size of the characteristic bottleneck was found to correlate with permeability predictions (Neumann et al. 2020). A short introduction to the definition and computation of this descriptor is given in Sect. S.7 of the Supplementary Information. In our paper samples, this descriptor shows a significant correlation with the absolute prediction error  $|v_{\text{CFD}} - v_{\text{PNM}}|$  in volume flow rates between CFD and PNM simulations, which suggests that it could carry meaningful information to further improve the volume flow rate prediction of PNM simulations. However, the descriptor also assumes only few discrete values, see Figure S5 of the Supplementary Information. This is likely because the size of the typical pore is close to the resolution of the image data. Higher resolved image data would be necessary to differentiate between microstructures that exhibit identical values for the size of the characteristic bottleneck at the current resolution. Thus, for the image data considered here, this descriptor is not suitable for the incorporation in a continuous prediction model, such as the linear regression models deployed in this study.

## 5 Structure–Property Relationships

In this section, we present six empirically derived regression models that will be used to predict the volume flow  $v_{\text{CFD}}$  as determined by CFD simulations, from geometric descriptors of pore space morphology introduced in Sect. 4, or from the volume flow prediction  $v_{\text{PNM}}$  of PNM simulations. In the following, the target variable of each regression will be denoted by  $v^{(i)}$  for  $i \in \{1, \dots, 6\}$ , which will be used to predict the value of  $v_{\text{CFD}}$ .

### 5.1 Regression Models for Structure–Property Relationships

The regression models that we consider are all of power-law type. Such models are easy to implement, and their predictions do not fundamentally differ from predictions using more complex relations (Hommel et al. 2018). The accompanying disregard of a critical porosity, often needed to describe media of small porosity (Hommel et al. 2018; Gebart 1992; Nabovati et al. 2009), is justified, as measured and simulated fluxes never vanished and the considered interval of porosity values is rather narrow, see Fig. 8a and b.

#### 5.1.1 Regression Models using only Geometric Descriptors

The simplest and most well-known relationship is given by

$$v^{(1)}(\varepsilon) = c_0 \varepsilon^{c_1}, \quad (17)$$

for some  $c_0, c_1 \in \mathbb{R}$ , where only the porosity  $\varepsilon$  is considered as an explanatory variable. This relation is widely used, as porosity is by far the most accessible geometric descriptor. With the next regression model, we add further geometric descriptors to capture the pore space morphology in more detail. It is given by

$$v^{(2)}(\varepsilon, \mu(\tau), \sigma(\tau), S) = c_0 \varepsilon^{c_1} \mu(\tau)^{c_2} \sigma(\tau)^{c_3} S^{c_4}, \quad (18)$$

for some  $c_0, c_1, c_2, c_3, c_4 \in \mathbb{R}$ . We can expect that  $v^{(1)}(\varepsilon) \neq v^{(2)}(\varepsilon, \mu(\tau), \sigma(\tau), S)$ , since the distribution of the lengths of transportation pathways through the material has a significant impact on the resulting volume flow rate. We also remark that the specific surface area  $S$  is not a dimensionless descriptor such as porosity or geodesic tortuosity. In order to analyze the improvement provided by the additional geometric descriptors considered in Eq. (18), it will be useful to additionally investigate the two simplified models

$$v^{(2,1)}(\varepsilon, \mu(\tau), \sigma(\tau)) = c_0 \varepsilon^{c_1} \mu(\tau)^{c_2} \sigma(\tau)^{c_3}, \quad (19)$$

for some  $c_0, c_1, c_2, c_3 \in \mathbb{R}$  and

$$v^{(2,2)}(\varepsilon, S) = c_0 \varepsilon^{c_1} S^{c_2}, \quad (20)$$

for some  $c_0, c_1, c_2 \in \mathbb{R}$ .

### 5.1.2 Regression Models Including the Results of PNM Simulations

Power laws in the form presented above are a helpful tool to test which descriptors indicate local intricacies of the pore space geometry that may have been oversimplified in the PNM simulations described in Sect. 3.2. On the other hand, we can assess the predictive power of the PNM itself with the relation

$$v^{(3)}(v_{\text{PNM}}) = c_0 v_{\text{PNM}}^{c_1}, \quad (21)$$

for some  $c_0, c_1 \in \mathbb{R}$ . This formulation allows for a comparison to  $v_{\text{CFD}}$  beyond strictly linear relations and automatically accounts for the difference in scale of the CFD and PNM predictions. By extending the relation given in Eq. (21) with geometric descriptors as additional factors, the improvement of the prediction depending on the considered descriptors indicates to which extend the descriptors "correct" the original PNM prediction. A combination of the models stated in Eqs. (17) and (21) is given by

$$v^{(4)}(\varepsilon, v_{\text{PNM}}) = c_0 \varepsilon^{c_1} v_{\text{PNM}}^{c_2}, \quad (22)$$

for some  $c_0, c_1, c_2 \in \mathbb{R}$ . Moreover, using the geometric descriptor  $r_{\text{max}}$ , we can further extend the relationship given in Eq. (22) and obtain

$$v^{(5)}(\varepsilon, r_{\text{max}}, v_{\text{PNM}}) = c_0 \varepsilon^{c_1} r_{\text{max}}^{c_2} v_{\text{PNM}}^{c_3}, \quad (23)$$

for some  $c_0, c_1, c_2, c_3 \in \mathbb{R}$ . Note that the median radius  $r_{\text{max}}$  and the specific surface area  $S$  have a physical unit, which introduces information on the typical length scale of the pore space. Finally, we consider a combination of Eqs. (18) and (21), which leads to

$$v^{(6)}(\varepsilon, \mu(\tau), \sigma(\tau), S, v_{\text{PNM}}) = c_0 \varepsilon^{c_1} \mu(\tau)^{c_2} \sigma(\tau)^{c_3} S^{c_4} v_{\text{PNM}}^{c_5}, \quad (24)$$

for some  $c_0, c_1, c_2, c_3, c_4, c_5 \in \mathbb{R}$ . The latter relationship contains five explanatory variables and, therefore, will likely provide the best fit simply by having the largest degree of freedom. Its purpose is primarily to serve as a reference when assessing the improvement of adding a particular descriptor.

## 5.2 Fitting and Validation of Regression Models

We briefly explain the fitting procedure for the models proposed in Sect. 5.1 and the criteria that we use to assess the goodness of fit. All models that we consider in the present study are power-type models of the general form

$$v = c_0 x_1^{c_1} \dots x_n^{c_n}, \quad (25)$$

for some  $n \in \{1, \dots, 5\}$ , where  $x_1, \dots, x_n \in \mathbb{R}$  are explanatory variables,  $c_0, \dots, c_n \in \mathbb{R}$  are coefficients that need to be determined, and  $v \in \mathbb{R}$  is the prediction of the volume flow determined by CFD simulations as described in Sect. 3.1. The explanatory variables  $x_1, \dots, x_n$  are either geometric descriptors as presented in Sect. 4 or the volume flow determined by PNM simulations as described in Sect. 3.2. Applying the natural logarithm to both sides of Eq. (25) yields

$$\log(v) = \log(c_0) + \sum_{i=1}^n c_i \log(x_i), \quad (26)$$

which is a linear regression problem with explanatory variables  $\log(x_1), \dots, \log(x_n)$ , response variable  $\log(v)$ , and coefficients  $\log(c_0), c_1, \dots, c_n \in \mathbb{R}$ . We determine the coefficients of this linear regression problem by means of the function `fitlm` in MATLAB (The MathWorks Inc 2022), which uses least square estimation. We then receive a prediction  $v$  of  $v_{\text{CFD}}$  by applying the exponential function to the prediction  $\log(v)$  obtained by Eq. (26).

In order to quantify the goodness of fit, we consider both the coefficient of determination (denoted by  $R^2$ ) and the mean absolute percentage error (denoted by MAPE). If  $m \in \mathbb{N} = \{1, 2, \dots\}$  is the number of data points available for the considered regression, these quantities are defined as

$$R^2 = 1 - \frac{\sum_{k=1}^m (\log(v_{\text{CFD},k}) - \log(v_k))^2}{\sum_{k=1}^m (\log(v_{\text{CFD},k}) - \bar{v}_{\text{CFD}})^2} \quad \text{and} \quad \text{MAPE} = \frac{100}{m} \sum_{k=1}^m \left| \frac{\log(v_{\text{CFD},k}) - \log(v_k)}{\log(v_{\text{CFD},k})} \right|, \quad (27)$$

where  $\log(v_{\text{CFD},1}), \dots, \log(v_{\text{CFD},m})$  are the (logarithmic) ground truth values resulting from CFD simulations as described in Sect. 3.1,  $\log(v_1), \dots, \log(v_m)$  are the corresponding predicted (logarithmic) values of the linear regression given in Eq. (26), and  $\bar{v}_{\text{CFD}}$  is the sample mean of  $\log(v_{\text{CFD},1}), \dots, \log(v_{\text{CFD},m})$ .

Note that the coefficient of determination  $R^2$  aims to quantify how much of the variance in the data is explained by the model fit, where, to ensure the proper interpretability of this quantity, it is crucial that the regression problem considered is linear (Spiess and Neumeyer 2010). Although we basically consider regressions of power type, the logarithmic transformation applied in Eq. (26) yields a linear regression problem. Therefore, we always use logarithmic values to evaluate both  $R^2$  and MAPE. That is, the ground truth values  $\log(v_{\text{CFD},1}), \dots, \log(v_{\text{CFD},m})$  are given by the logarithms of the simulated volume flows, computed as described in Sect. 3.1, and the predictions  $\log(v_1), \dots, \log(v_m)$  are the predicted volume flows obtained from the linear regression problem described by Eq. (26). In this way, we ensure that the resulting values allow for an adequate comparison between the regression models stated in Sect. 5.1. However, due to the limited amount of available data, we always use the same data to fit the coefficients  $\log(c_0), c_1, \dots, c_n \in \mathbb{R}$  in Eq. (26) as we do to evaluate the statistics  $R^2$  and MAPE defined in Eq. (27).

## 6 Results and Discussion

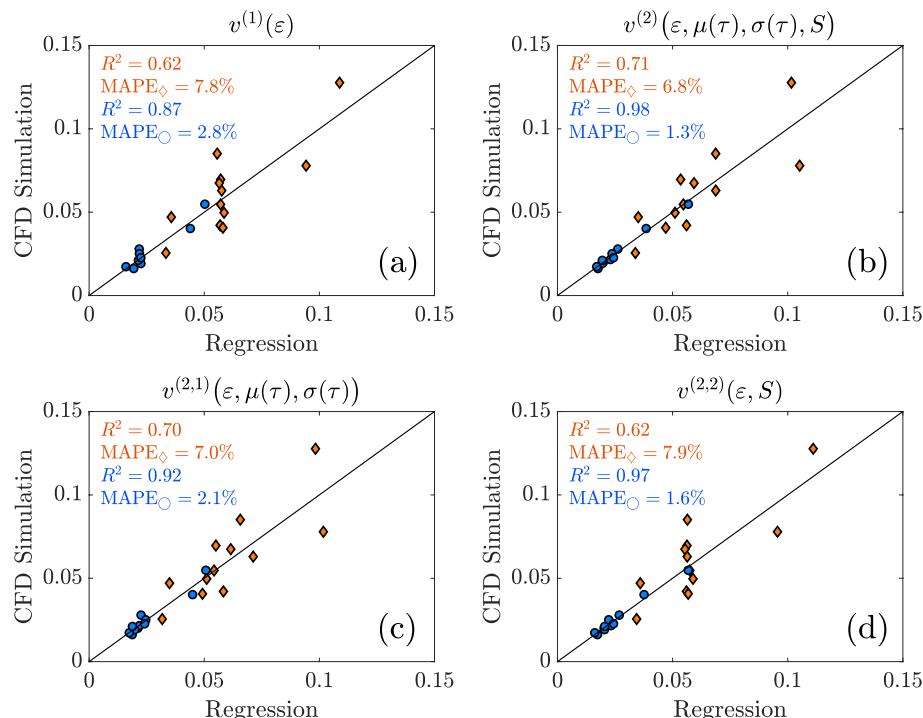
We now present the results we obtained for the fitted regression models stated in Sect. 5.1. At first, all regression models are fitted separately to both data sets of uncompressed and compressed paper sheets, yielding two separate sets of coefficients for each regression model. In this way, we can analyze the predictive power of the involved descriptors in a sample-specific manner, yielding insights into how the dependency structure between geometric descriptors and the volume flow rates obtained by CFD and PNM simulations changes between different samples.

In Sect. 6.1, we consider regression models that involve only purely geometric descriptors of pore space. Afterward, in Sect. 6.2, we show how incorporating the results of PNM

simulations can be used to improve the predictions of the volume flow rate. Finally, in Sect. 6.3, we analyze how well the considered regression models generalize across uncompressed and compressed paper sheets, by using only one set of coefficients for each regression model that has been jointly determined from both data sets. The analysis of these regression models, and the comparison to the separately fitted regression models discussed before, will provide insight into how to formulate general relationships that apply to varying grades of paper sheets with different morphologies. The values determined for each of the coefficients  $c_0, \dots, c_n$  of the regression models considered are listed in Tables S1 – S3 of the Supplementary Material.

## 6.1 Regression Models Using Only Geometric Descriptors

In this section, we only consider the purely geometric regression models stated in Sect. 5.1.1. These will give insights into predicting volume flow rates by CFD simulations based only on geometric information of the pore space morphology. The results of the fits are visualized in Fig. 9, where the predictions of the corresponding regression models are plotted against the flow rates of CFD simulations.



**Fig. 9** Scatter plots of volume flow rates obtained by CFD simulations versus the values predicted by the respective regression models. Separate fits of regression models have been determined for the data points of the uncompressed (orange diamonds) and compressed sample (blue circles). The coefficient of determination  $R^2$  and the mean absolute percentage MAPE are displayed for both samples on each plot. As a guide to the eye, the black line highlights the diagonal on which the results of CFD simulation and regression coincide

As mentioned above, for each of the regression models, two sets of coefficients  $c_0, \dots, c_n$  have been determined separately for uncompressed and compressed paper sheets. Therefore, Fig. 9 also shows two values for the coefficient of determination  $R^2$  and the mean absolute percentage error MAPE, displayed in orange and blue for the uncompressed and compressed sample, respectively.

We start by considering the simplest model  $v^{(1)}$  given in Eq. (17), which attempts to establish a direct connection between porosity  $\varepsilon$  and volume flow rate  $v_{\text{CFD}}$  obtained by CFD simulations, see Fig. 9a. While  $v^{(1)}$  predicts the overall trend quite well, it cannot distinguish samples with a similar porosity, which results in clusters of vertically arranged data points in Fig. 9a. This effect results in the low value of  $R^2 = 0.62$  in the uncompressed case, but is less detrimental in the compressed sample, where we receive a value of  $R^2 = 0.87$ .

The vertically clustered points in Fig. 9a clearly indicate that there are factors influencing the volume flow that cannot be explained by porosity alone. This is consistent with the situation found in similar paper sheets (Leitl et al. 2023). Therefore, our next step is to consider three further descriptors of pore space morphology that are all incorporated in the model  $v^{(2)}$  given in Eq. (18). The first two descriptors relate to the distribution of geodesic tortuosity  $\tau$ , which enters  $v^{(2)}$  in the form of its mean value  $\mu(\tau)$  and standard deviation  $\sigma(\tau)$ , while the third is the specific surface area  $S$ . The fit with respect to  $v^{(2)}$  substantially improves the fit by  $v^{(1)}$  regardless of the sample, see Fig. 9b. In particular, for the compressed sample a value of  $R^2 = 0.98$  and for the uncompressed sample, a value of  $R^2 = 0.71$  is achieved.

Obviously, it is not surprising that we achieve an improved fit in both cases when refining the model through additional descriptors. However, individual descriptors improve the fit by varying degrees in a sample-specific way. To make this clear, we consider the regression models  $v^{(2,1)}$  and  $v^{(2,2)}$  given in Eqs. (19) and (20), which both use a subset of the descriptors considered in  $v^{(2)}$ , see Fig. 9c and d. For the uncompressed sample (orange diamonds in Fig. 9), the accuracy of  $v^{(2,1)}$  is almost identical to that of  $v^{(2)}$ , with a coefficient of determination of  $R^2 = 0.70$  for  $v^{(2,1)}$ , compared to  $R^2 = 0.71$  for  $v^{(2)}$ . Thus, for the data set of this sample, one can argue that information on the specific surface area  $S$  is redundant to predict the volume flow rate  $v_{\text{CFD}}$  obtained by CFD simulations, while information on the geodesic tortuosity is more relevant. On the other hand, for the compressed sample, the regression model  $v^{(2,2)}$  yields a value of  $R^2 = 0.97$  in comparison with a value of  $R^2 = 0.98$  for the more complex model  $v^{(2)}$ . Thus, in this case, information on the specific surface area  $S$  provides an enormous improvement compared to the regression model  $v^{(1)}$ , while the information on geodesic tortuosity seems redundant.

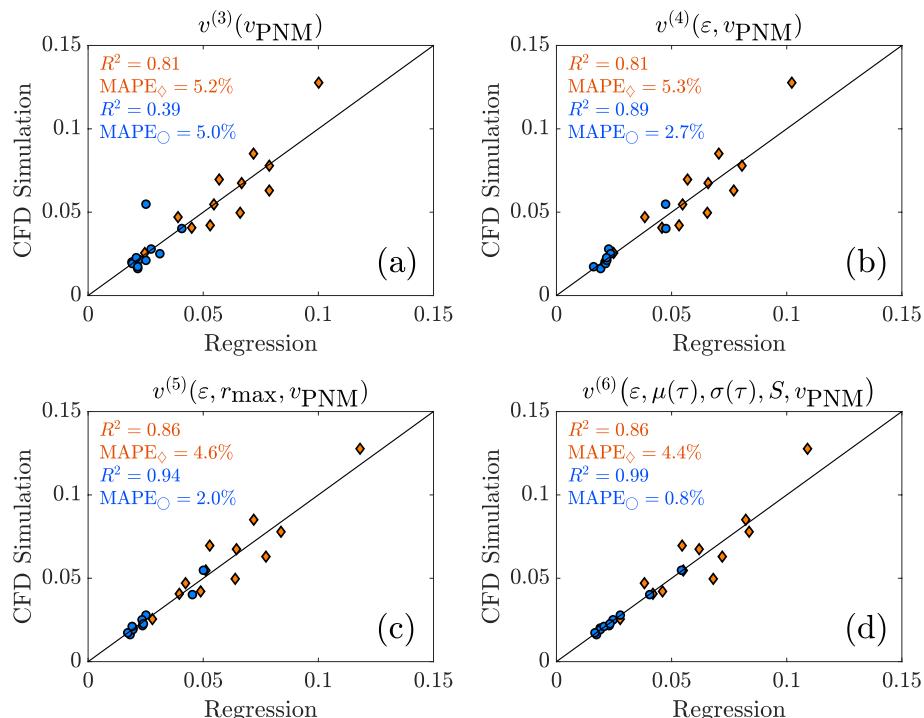
Intuitively, one might assume that any new descriptor, which is highly correlated with already present descriptors, would not provide enough new information and therefore cannot significantly improve the fit. However, as shown in Fig. 8a, the porosity  $\varepsilon$  and the specific surface area  $S$  are significantly correlated in the compressed case, with only a low correlation coefficient in the uncompressed case. Nevertheless, the compressed case benefits more from knowledge on the specific surface area  $S$ , while the uncompressed case sees more improvement by incorporating information on the geodesic tortuosity.

This illustrates that a descriptor cannot be deemed as redundant or not, just based on its correlations with other (already present) descriptors. Note that the presence of correlations among explanatory variables affects the variance and interpretability of the values of the estimated coefficients  $c_0, \dots, c_n$  provided in Tables S1 – S3. In particular, we cannot formulate a quantitative statement that a certain increase of an explanatory variable would have a certain effect on the predictor variable, as it might not be possible to increase an

explanatory variable without affecting the others. However, these correlations do in general not influence the quality of the obtained fit (Kutner et al. 2004). In turn, the absence of correlations among explanatory variables does not generally guarantee an improved fit. In many machine learning approaches, feature selection is performed according to a so-called maximum-relevance-minimum-redundancy criterion, which aims to weigh the correlation between a given feature and the target observation against the correlation between the given feature and other, already present, features (Peng et al. 2005; van der Linden et al. 2016). However, such a systematic approach requires an accurate estimation of the probability distributions of involved features, which is not feasible in our case due to the small data base. Instead, we follow the same idea by manually evaluating our metrics for the goodness of fit and analyzing the correlation between the present descriptors.

## 6.2 Combining Geometric Information with PNM Simulations

We now consider the regression models  $v^{(3)}$  to  $v^{(6)}$  stated in Sect. 5.1.2, which all involve the volume flow rate  $v_{\text{PNM}}$  as determined by PNM simulations. The model  $v^{(3)}$  uses only  $v_{\text{PNM}}$ , while  $v^{(4)}$ ,  $v^{(5)}$ , and  $v^{(6)}$  combine information from geometric descriptors and PNM simulations. The analysis of these models provides insight about the accuracy



**Fig. 10** Scatter plots of volume flow rates obtained by CFD simulations versus the values predicted by regression models  $v^{(3)}$  to  $v^{(6)}$  (panels 10a to 10d) aiming at correcting the flow rate predictions obtained by PNM simulations. Separate fits of the regression model to the uncompressed (orange diamonds) and compressed sample (blue circles). The coefficient of determination  $R^2$  and the mean absolute percentage MAPE are displayed for both samples on each plot. As a guide to the eye, the black line highlights the diagonal on which the results of CFD simulation and regression coincide

of predicting the results of CFD simulations based on those of PNM simulations, and whether it is possible to find geometric descriptors of pore space morphology that mitigate the discrepancies between the two predictions. Similarly to the previous section, all the models considered use two sets of coefficients  $c_0, \dots, c_n$ , which have been estimated separately based on the data of uncompressed and compressed paper sheets, respectively.

Figure 10a shows how well flow rates obtained from PNM simulations predict the CFD flow values, using the regression model  $v^{(3)}$ . The overall fit is clearly better for the uncompressed case than for the compressed case, with a higher  $R^2$ -value and lower MAPE. However, the lack of accuracy in the compressed case ( $R^2 = 0.39$ ) is mainly due to an outlier associated with a CFD volume flow of  $v_{\text{CFD}} = 0.05 \text{ m/s}$ , see Fig. 10a.

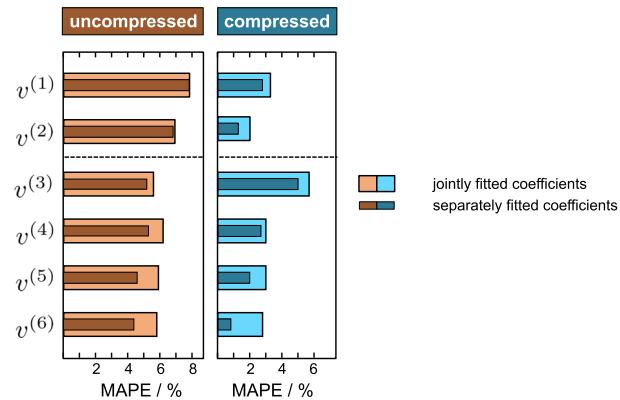
The regression model  $v^{(4)}$  considers porosity as an additional descriptor that can correct flow rates determined by PNM simulations. With the additional descriptor, the model can correct for the above mentioned outlier, increasing the coefficient of determination from  $R^2 = 0.39$  to  $R^2 = 0.89$  in the compressed case, see Fig. 10b. Furthermore, a comparison between the fits of  $v^{(4)}$  and  $v^{(1)}$  also shows that incorporating PNM simulations already significantly improves the classical approach of predicting volume flow rates only by means of porosity.

The regression models  $v^{(5)}$  and  $v^{(6)}$  reveal, in analogy to the transition from  $v^{(1)}$  to  $v^{(2)}$ , which geometric descriptors in addition to  $\varepsilon$  significantly correct the PNM-predicted flow rates and whether a strong correlation of a descriptor with porosity (or the absence thereof) controls the extent of improvement. Recall that  $v^{(5)}$  considers the porosity  $\varepsilon$  and the median pore radius  $r_{\text{max}}$ , obtained from the continuous pore size distribution. Here we observe a similar phenomenon as in the previous section. The median pore radius  $r_{\text{max}}$  is strongly correlated with porosity  $\varepsilon$  for both samples, see Fig. 8d. Despite this strong correlation, the model provided by  $v^{(5)}$  fits the data of the CFD volume flows significantly better than  $v^{(4)}$  for both samples, see Fig. 10c. Finally,  $v^{(6)}$  considers all geometric descriptors that were already used in  $v^{(2)}$  as explanatory variables, in addition to the PNM-predicted flow rate. In comparison with  $v^{(2)}$ , the additional information of  $v_{\text{PNM}}$  still leads to an improvement of an already highly accurate prediction, i.e., from an  $R^2$ -value of 0.98 to 0.99, see Fig. 10d. Most notably, the models  $v^{(5)}$  and  $v^{(6)}$  yield nearly identical precisions in the uncompressed case, although  $v^{(6)}$  uses the three descriptors  $\mu(\tau)$ ,  $\sigma(\tau)$ , and  $S$  to replace only the one descriptor  $r_{\text{max}}$  in  $v^{(5)}$ . Yet again, this is in contradiction to our expectations, given that the porosity  $\varepsilon$  and the median pore radius  $r_{\text{max}}$  have a high correlation coefficient of  $\rho = 0.89$  in the uncompressed case.

### 6.3 Generalizing Relationships Across Different Paper Grades

With the regression models at hand, the question arises whether the models would also satisfactorily predict the CFD volume flow rates for a wider range of microstructures. To test this, we now determine a single set of coefficients for each of the regression models stated in Sect. 5.1, which is jointly determined from all data points of both the combined uncompressed and compressed samples. As we want to compare the performance of these jointly fitted models to the separately fitted models of the previous sections, we also require the metric for the goodness of fit to be comparable. Therefore, we again determine the values of the coefficient of determination  $R^2$  and the mean absolute percentage error MAPE for each regression. More precisely, we evaluate these metrics once only on the data points of the uncompressed sample and once only on the compressed sample. In this way, we obtain

**Fig. 11** Mean average percentage error (MAPE) between CFD-predicted volume flow rates and predictions of the regression models  $v^{(1)}, \dots, v^{(6)}$ , each fitted to the combined data points of uncompressed and compressed paper (wide bars), or separately to uncompressed paper and compressed paper, respectively (narrow bars)



a value of  $R^2$  and MAPE for each sample and each of the jointly fitted models that are directly comparable to the  $R^2$ - and MAPE-values of the separately fitted models.

Figure 11 collects the MAPE-values for all jointly fitted and separately fitted models, evaluated for both data sets of the uncompressed and the compressed sample. Bars with a light shading indicate the MAPE-values of the jointly fitted model, while the narrow bars with darker shading show the MAPE-values from the previous sections, in which separate sets of coefficients  $c_0, \dots, c_n$  were used for each sample. Corresponding scatter plots for these regression models are shown in Figure S1 of the Supplementary Information.

The models  $v^{(1)}, \dots, v^{(6)}$  considered in Fig. 11 are ordered with increasing complexity (from top to bottom): Regression models  $v^{(1)}, v^{(2)}$  exclusively contain geometric descriptors of the pore space, while regression models  $v^{(3)}, \dots, v^{(6)}$  are associated with comparison and corrections of the predictions by means of PNM simulations. Obviously, with increasing complexity of the models, the values of MAPE are decreasing. However, we can see that in some cases the added complexity leads only to an improved model if it can specialize to certain types of microstructures. For example, this is the case for  $v^{(6)}$  in the compressed case, see Fig. 11. In general, the more complex the model, the greater the discrepancy between the separately fitted models and the jointly fitted model. In particular, we see that for both samples, the jointly fitted models  $v^{(4)}, \dots, v^{(6)}$  all show very similar performance, indicating that the additional complexity of  $v^{(5)}$  and  $v^{(6)}$  can only be capitalized on when fitting to specific data sets, but not when generalizing across different grades of paper. However, the decrease in MAPE between  $v^{(1)}$  and  $v^{(4)}$  is consistent between both samples, indicating that combining the results of the PNM simulations with geometric descriptors yields an improvement regardless of the specific paper grade considered.

In general, using porosity  $\varepsilon$  as a correction factor in model  $v^{(4)}$  for the prediction obtained by PNM simulations helps to better explain the flow rates in the compressed sample, i.e., the denser sample. The descriptor  $r_{\max}$  also has a significant impact for the compressed sample. However, it remains unclear whether this is due to sample-specific details of the pore network model. To decide this, the structures harboring local flows need to be resolved in more detail. In such a local analysis, the structure of the pore network model could also play a role, as PNM simulations on conduits of the same shape tend to neglect too many details of the actual shape of the interface formed between connected pores (Zhao et al. 2020).

## 7 Summary and Conclusion

In this study, the volume fluxes of air through various cutouts of two different samples of porous paper sheets were determined by using CFD and PNM simulations. The fluxes from CFD simulations serve as ground truth in our study, while PNM simulations provide a simplified and computationally much less demanding alternative. In addition to volume fluxes, various geometric descriptors of the pore space have been determined for each of the cutouts based on the tomographic  $\mu$ -CT image data. In particular, we focused on analyzing the performance of the PNM simulations in comparison with CFD simulations and investigating whether additional information on the pore space morphology through geometric descriptors can be used to improve the fluxes predicted by PNM simulations. We also directly analyzed structure–property relationships between geometric descriptors and volume fluxes of CFD simulations. This was done by fitting six different regression models of power-law type to the data obtained for the cutouts of both samples. These regression models were fitted separately to the two data sets of compressed and uncompressed paper samples to analyze the performance of the models in a sample-dependent way. Subsequently, we also fitted the models to the joint data set of both samples to see how well the individual models generalize across both samples.

Our analysis showed that, while volume fluxes determined by PNM simulations qualitatively resemble the fluxes determined by CFD simulations, involving additional geometric descriptors significantly improves the accuracy of the predictions. From the sample-specific analysis, we see that the ideal choice of geometric descriptors depends on the chosen sample, which is in line with previous results showing altered correlation structures of geometric descriptors in paper sheets after compression (Neumann et al. 2024). However, even simply including porosity in the regression model already yields an improvement for both samples. In particular, the regression model using only porosity and PNM fluxes in some cases outperforms more complex relationships that directly model the CFD fluxes by multiple geometric descriptors of pore space, which underlines the potential of PNM as a more cost-efficient simulation tool.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11242-025-02250-9>.

**Author Contributions** E.M.C., R.S., A.H., and I.M. acquired the micro-CT scans. M.N., E.M.C., and A.H. segmented the image data. U.H. suggested use cases and supported the interpretation of the model predictions. M.N. and P.G. performed the statistical analysis of the microstructure data. A.R. and P.L. performed the CFD simulations. E.B. and M.F. performed the PNM simulations. M.N., P.G., and K.Z. conceptualized and wrote the manuscript. M.N., P.G., V.S., and K.Z. were involved in the editing of the manuscript. All authors have read and approved the final version of the manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL. The authors appreciate financial support from the Christian Doppler Research Association, the Federal Ministry for Digital and Economic Affairs, and the National Foundation for Research, Technology, and Development, Austria. M.N. acknowledges funding by the German Research Foundation (DFG) under Project ID 390874152 (POLiS Cluster of Excellence, EXC 2154). K.Z. thanks the "TU Graz Lead Project LP-03: Porous Materials @ Work for Sustainability" for inspiration.

**Data Availability** Additional data are available as supplementary information. The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors have no relevant financial or non-financial interests to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

Ademmer, M., Su, P.-H., Dodell, L., Asenbauer, J., Osenberg, M., Hilger, A., Chang, J.-K., Manke, I., Neumann, M., Schmidt, V., Bresser, D.: Unveiling the impact of cross-linking redox-active polymers on their electrochemical behavior by 3D imaging and statistical microstructure analysis. *J. Physi. Chem. C* **127**, 19366–19377 (2023)

Akbari, M., Sinton, D., Bahrami, M.: Viscous flow in variable cross-section microchannels of arbitrary shapes. *Int. J. Heat Mass Transf.* **54**, 3970–3978 (2011)

Arganda-Carreras, I., Kaynig, V., Rueden, C., Eliceiri, K.W., Schindelin, J., Cardona, A., Seung, H.S.: Trainable Weka segmentation: a machine learning tool for microscopy pixel classification. *Bioinformatics* **33**, 2424–2426 (2017)

Aslannejad, H., Hassanzadeh, S.M., Raoof, A., de Winter, D.A.M., Tomozeiu, N., van Genuchten, M.T.: Characterizing the hydraulic properties of paper coating layer using FIB-SEM tomography and 3D pore-scale modeling. *Chem. Eng. Sci.* **160**, 275–280 (2017)

Barman, S., Rootzén, H., Bolin, D.: Prediction of diffusive transport through polymer films from characteristics of the pore geometry. *AIChE J.* **65**, 446–457 (2019)

Bernabé, Y.: Gaseous flow through heterogeneous, partially connected networks of pipes. *Sci. Rep.* **8**, 14956 (2018)

Blunt, M.J., Bijeljic, B., Dong, H., Gharbi, O., Iglauder, S., Mostaghimi, P., Paluszny, A., Pentland, C.: Pore-scale imaging and modelling. *Adv. Water Resour.* **51**, 197–216 (2013)

Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J.: *Stochastic Geometry and its Applications*. J. Wiley & Sons, 3rd edition (2013)

Chung, T.J.: *Computational Fluid Dynamics*. Cambridge University Press (2002)

Dijkstra, E.W.: A note on two problems in connexion with graphs. *Numer. Math.* **1**, 269–271 (1959)

Dodell, L., Neumann, M., Osenberg, M., Hilger, A., Studer, G., Esser, B., Manke, I., Schmidt, V.: Quantifying local heterogeneities in the 3D morphology of X-PVMPt battery electrodes based on FIB-SEM measurements. *ACS Mater. Lett.* **7**, 1914–1921 (2025)

Fischer, R., Schlepütz, C.M., Hegemann, D., Rossi, R.M., Derome, D., Carmeliet, J.: Four-dimensional imaging and free-energy analysis of sudden pore-filling events in wicking of yarns. *Phys. Rev. E* **103**, 053101 (2021)

Gebart, R.: Permeability of unidirectional reinforcements for RTM. *J. Compos. Mater.* **26**, 1100–1133 (1992)

Gombosi, T.I.: *Gaskinetic Theory*. Cambridge University Press, Cambridge Atmospheric and Space Science Series (1994)

Gostick, J.T.: Versatile and efficient pore network extraction method using marker-based watershed segmentation. *Phys. Rev. E* **96**, 023307 (2017)

Gostick, J., Aghighi, M., Hinebaugh, J., Tranter, T., Hoeh, M.A., Day, H., Spellacy, B., Sharqawy, M.H., Bazylak, A., Burns, A., et al.: OpenPNM: a pore network modeling package. *Comput. Sci. Eng.* **18**, 60–74 (2016)

Gostick, J., Khan, Z., Tranter, T., Kok, M., Agnaou, M., Sadeghi, M., Jervis, R.: PoreSpy: a python toolkit for quantitative analysis of porous media images. *J. Open Source Software* **4**, 1296 (2019)

Holzer, L., Iwanschitz, B., Hocker, T., Keller, L., Pecho, O.M., Sartoris, G., Gasser, P., Münch, B.: Redox cycling of Ni-YSZ anodes for solid oxide fuel cells: influence of tortuosity, constriction and percolation factors on the effective transport properties. *J. Power Sources* **242**, 179–194 (2013)

Holzer, L., Wiedenmann, D., Münch, B., Keller, L., Prestat, M., Gasser, P., Robertson, I., Grobety, B.: The influence of constrictivity on the effective transport properties of porous layers in electrolysis and fuel cells. *J. Mater. Sci.* **48**, 2934–2952 (2013)

Holzer, L., Marmet, P., Fingerle-Straß, M., Wiegmann, A., Neumann, M., Schmidt, V.: *Tortuosity and Microstructure Effects in Porous Media: Classical Theories. Springer, Empirical Data and Modern Methods* (2023)

Hommel, J., Coltman, E., Class, H.: Porosity–permeability relations for evolving pore space: A review with a focus on (bio-)geochemically altered porous media. *Transp. Porous Media* **124**, 589–629 (2018)

Hutten, I.M.: *Handbook of Nonwoven Filter Media*. Butterworth-Heinemann (2015)

Ioannidis, M., Chatzis, I.: Network modeling of pore structure and transport-properties of porous-media. *Chem. Eng. Sci.* **48**, 951 (1993)

ISO 5636-5:2013. Paper and board: Determination of air permeance (medium range) – Part 5: Gurley method. Standard, Geneva: Standard, International Organization for Standardization (2013)

Jeulin, D.: *Morphological Models of Random Structures*. Springer (2021)

Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W.: *Applied Linear Statistical Models*. McGraw-Hill, 5th edition (2004)

Leal, L.G.: *Advanced Transport Phenomena: Fluid Mechanics and Convective Transport Processes*. Cambridge University Press (2007)

Leitl, P.: Modelling the air flow through a paper web using computational fluid dynamics. Master's thesis, Graz University of Technology (2020)

Leitl, P., Machado-Charry, E., Baikova, E., Neumann, M., Hirn, U., Schmidt, V., Zojer, K.: Joint distributions of local pore space properties quantitatively explain simulated air flow variations in paper. *Transp. Porous Media* **148**, 1–22 (2023)

Machado Charry, E., Neumann, M., Lahti, J., Schennach, R., Schmidt, V., Zojer, K.: Pore space extraction and characterization of sack paper using  $\mu$ -CT. *J. Microscopy* **272**, 35–46 (2018)

Maurer, C.R., Qi, R., Raghavan, V.: A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 265–270 (2003)

Miao, X., Gerke, K.M., Sizonenko, T.O.: A new way to parameterize hydraulic conductances of pore elements: a step towards creating pore-networks without pore shape simplifications. *Adv. Water Resour.* **105**, 162–172 (2017)

Morgan, D., Jacobs, R.: Opportunities and challenges for machine learning in materials science. *Annual Rev. Mater. Res.* (2020)

Nabovati, A., Llewellyn, E.W., Sousa, A.C.M.: A general model for the permeability of fibrous porous media based on fluid flow simulations using the lattice Boltzmann method. *Compos. A Appl. Sci. Manuf.* **40**, 860–869 (2009)

Neumann, M., Stenzel, O., Willot, F., Holzer, L., Schmidt, V.: Quantifying the influence of microstructure on effective conductivity and permeability: virtual materials testing. *Int. J. Solid Struct.* **184**, 211–220 (2020)

Neumann, M., Machado-Charry, E., Baikova, E., Hilger, A., Hirn, U., Schennach, R., Manke, I., Schmidt, V., Zojer, K.: Capturing centimeter-scale local variations in paper pore space via  $\mu$ -CT: a benchmark study using calendered paper. *Microsc. Microanal.* **27**, 1305–1315 (2021)

Neumann, M., Machado Charry, E., Zojer, K., Schmidt, V.: On variability and interdependence of local porosity and local tortuosity in porous materials: a case study for sack paper. *Methodol. Comput. Appl. Probability* **23**, 613–627 (2021)

Neumann, M., Ademmer, M., Osenberg, M., Hilger, A., Wilde, F., Muench, S., Hager, M.D., Schubert, U.S., Manke, I., Schmidt, V.: 3D microstructure characterization of polymer battery electrodes by statistical image analysis based on synchrotron X-ray tomography. *J. Power Sources* **542**, 231783 (2022)

Neumann, M., Gräfensteiner, P., Machado Charry, E., Hirn, U., Hilger, A., Manke, I., Schennach, R., Schmidt, V., Zojer, K.: R-vine copulas for data-driven quantification of descriptor relationships in porous materials. *Adv. Theory Simul.* **7**, 2301261 (2024)

Oh, W., Lindquist, B.: Image thresholding by indicator kriging. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**, 590–602 (1999)

Ohser, J., Schladitz, K.: *3D Images of Materials Structures: Processing and Analysis*. Wiley-VCH (2009)

Peng, H., Long, F., Ding, C.: Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**, 1226–1238 (2005)

Rodriguez de Castro, A., Agnau, M., Gostick, J.: Predicting shear-thinning fluid flows in porous media using pore network modeling: simulations and experimental validation. *Transp. Porous Media* **149**, 453–478 (2023)

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A.: Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–82 (2012)

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.W., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A.: Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676 (2012)

Schladitz, K., Ohser, J., Nagel, W.: Measuring intrinsic volumes in digital 3D images. In A. Kuba, L. G. Nyúl, and K. Palágyi, editors, *International Conference on Discrete Geometry for Computer Imagery (DGCI)*, pages 247–258, Berlin (2006)

Serra, J.: *Image Analysis and Mathematical Morphology*. Academic Press (1982)

Sochi, T.: Newtonian flow in converging-diverging capillaries. *Int. J. Model. Simul. Sci. Comput.* **04**, 1350011 (2013)

Sohr, H.: *The Navier-Stokes Equations: An Elementary Functional Analytic Approach*. Springer (2012)

Soille, P.: *Morphological Image Analysis: Principles and Applications*. Springer, 2nd edition (2003)

Spiess, A.-N., Neumeyer, N.: An evaluation of  $R^2$  as an inadequate measure for nonlinear models in pharmaceutical and biochemical research: a Monte Carlo approach. *BMC Pharmacol.* **10**, 6 (2010)

Sternberg, S.R.: Biomedical image processing. *Computer* **16**, 22–34 (1983)

Succi, S., Benzi, R., Higuera, F.: The lattice Boltzmann equation: a new tool for computational fluid-dynamics. *Physica D* **47**, 219–230 (1991)

The MathWorks Inc. Matlab, (2022). Version: 9.12 (R2022a)

Torquato, S.: *Random Heterogeneous Materials: Microstructure and Macroscopic Properties*. Springer (2022)

Valvatne, P.H., Blunt, M.J.: Predictive pore-scale modeling of two-phase flow in mixed wet media. *Water Resour. Res.* **40**, W07406 (2004)

van der Linden, J.H., Narsilio, G.A., Tordesillas, A.: Machine learning framework for analysis of transport through complex networks in porous, granular media: A focus on permeability. *Physical Review E* (2016)

Whitaker, S.: Flow in porous media I: a theoretical derivation of Darcy's law. *Transp. Porous Media* **1**, 3–25 (1986)

Xu, K., Wei, W., Chen, Y., Tian, H., Xu, S., Cai, J.: A pore network approach to study throat size effect on the permeability of reconstructed porous media. *Water* **14**, 77 (2022)

Yang, X., Mehmani, Y., Perkins, W.A., Pasquali, A., Schönherr, M., Kim, K., Perego, M., Parks, M.L., Trask, N., Balhoff, M.T., Richmond, M.C., Geier, M., Krafczyk, M., Luo, L.-S., Tartakovsky, A.M., Scheibe, T.D.: Intercomparison of 3D pore-scale flow and solute transport simulation methods. *Adv. Water Resour.* **95**, 176–189 (2016)

Zhao, J., Qin, F., Derome, D., Kang, Q., Carmeliet, J.: Improved pore network models to simulate single-phase flow in porous media by coupling with lattice Boltzmann method. *Adv. Water Resour.* **145**, 103738 (2020)

Zhao, J., Qin, F., Kang, Q., Qin, C., Derome, D., Carmeliet, J.: A dynamic pore network model for imbibition simulation considering corner film flow. *Water Resources Res.* **58**, ER032332 (2022)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.