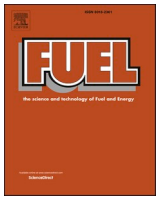
Fuel 321 (2022) 124044

Contents lists available at ScienceDirect

Fuel

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



Full Length Article

Machine learning modeling of permeability in 3D heterogeneous porous media using a novel stochastic pore-scale simulation approach

Olubukola Ishola, Javier Vilcáez *

Boone Pickens School of Geology, Oklahoma State University, Stillwater, OK 74078, USA

ARTICLE INFO

Keywords:
Permeability
Pore-scale modeling
Stochastic simulations
Machine learning

ABSTRACT

Accurate predictions of rock permeability is critical for resource exploration and environmental management. To improve on existing approaches to permeability prediction, this study employed a stochastic pore-scale simulation approach. The petrophysical properties needed for the implementation of this approach are porosity and pore size distribution (PSD) of rock samples which can be obtained easily from mercury injection capillary pressure measurements. The approach was tested on four carbonate and five siliciclastic rock cores. To consider a wide range of possible pore connectivity scenarios that can be associated to the same PSD and porosity, the employed stochastic pore-scale simulation approach involves the generation of hundreds of 3D pore microstructures of the same PSD and porosity but different stochastic pore connectivity. Permeability is calculated by averaging the permeability distribution obtained from pore-scale flow simulations through the generated 3D pore microstructures. Permeability estimations were closer to measured permeability with this approach than with five deterministic empirical model equations. Machine learning was used to reduce the required number of pore-scale simulations by 157 times and reproduced permeability estimated from pore-scale flow simulations with a mean absolute percentage error of 10%.

1. Introduction

Predictions of flow properties in heterogeneous porous media are highly problematic. In general, assessments of subsurface flow and reactive transport processes through rocks requires knowledge of permeability [1–3]. Permeability is the ease of fluid flow through porous media [4,5]. Fundamentally, permeability is a function of the pore microstructure of porous media [6–8]. True permeability of rocks is obtained through laboratory measurement from core plugs (rock samples) and pressure transient analysis of well test data [9]. In absence of true permeability measurements, permeability is traditionally estimated from model equations that relate permeability to other petrophysical rock properties. Kozeny [10,11] proposed one of the foremost equations used in permeability estimations. The Kozeny equation is formed from the combination of Darcy's Law [12], the Hagen-Poiseuille velocity equation [13], and the concept of tortuosity [14]. Kozeny's equation is expressed mathematically as [11]:

$$K = \frac{\phi d^2}{\tau} \tag{1}$$

where K is the permeability, τ is the tortuosity defined as the ratio of actual length of flow path in the porous media to the length of flow path in the absence of porous media, d is the pore hydraulic diameter, and ϕ is the porosity of the porous media. The equation was later modified into the Kozeny-Carman equation [11,15] The Kozeny-Carman equation is given by [11]:

$$K = \frac{\phi d^2}{16\ K_k} \tag{2}$$

where K_k is the Kozeny constant which captures the effect of tortuosity, particle shape, and connectivity of pores.

The dependence of the Kozeny and Kozeny-Carman equations on tortuosity [11] implies that it is not suited for making permeability prediction in heterogenous porous media. This is due to the deterioration of permeability-tortuosity at increased levels of heterogeneity [16]. Also, porosity, which is the key parameter in Kozeny-Carman equation [11,15] is known to have poor correlation with permeability [17]. Permeability have been shown to vary up to five orders of magnitude for a very narrow porosity range [18].

Pore size distribution of rocks is deemed to provide better estimation

* Corresponding author.
E-mail address: vilcaez@okstate.edu (J. Vilcáez).

https://doi.org/10.1016/j.fuel.2022.124044
Received 13 February 2022; Received in revised form 26 March 2022; Accepted 29 March 2022
Available online 2 April 2022

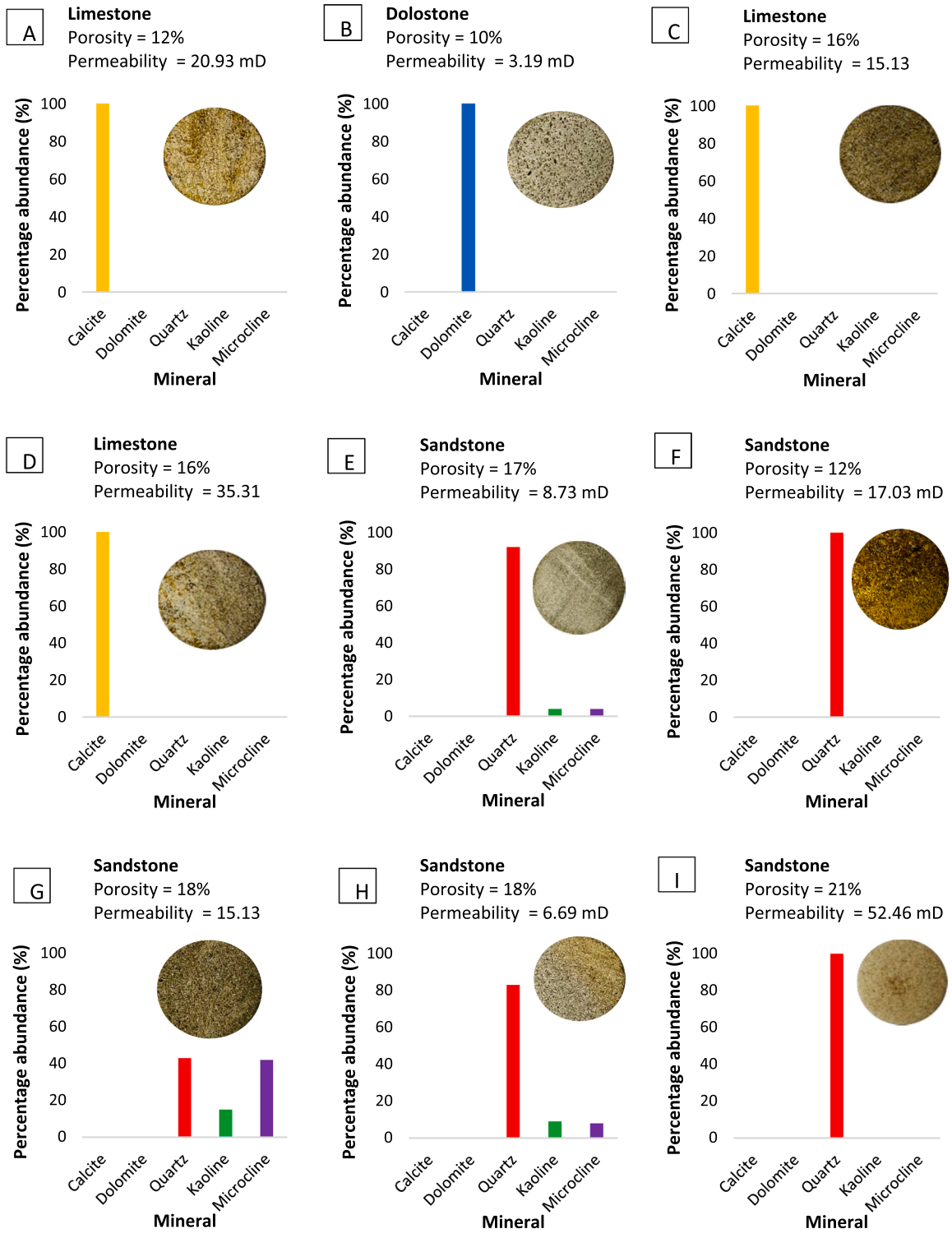


Fig. 1. Mineralogical composition of the core plugs employed in this study and their measured petrophysical properties.

of permeability when combined with porosity [18]. This can be implemented using Nuclear Magnetic Resonance (NMR) T2 data through the Timur-Coates (TC) equation [19] and the Schlumberger-Doll-Research (SDR) equation [20]. Westphal [18] identified heterogeneity as a limitation in the application of TC and SDR equations in carbonate rocks because they assume even distribution of pore connectivity and employ total porosity instead of effective porosity in their respective equations. Mercury injection capillary pressure (MICP) is another routine approach used to make permeability predictions using empirical models. Several authors have used percentage mercury saturation or geometric average of pore throat sizes as a parameter to predict permeability [22–28]. The use of MICP for permeability predictions is empirical and it depends heavily on the data distribution that are used in deriving the respective model equations. In general, empirical model equations are constrained by the data used in their derivations. This is a challenge for heterogeneous rocks such as carbonate rocks because pore connectivity in these

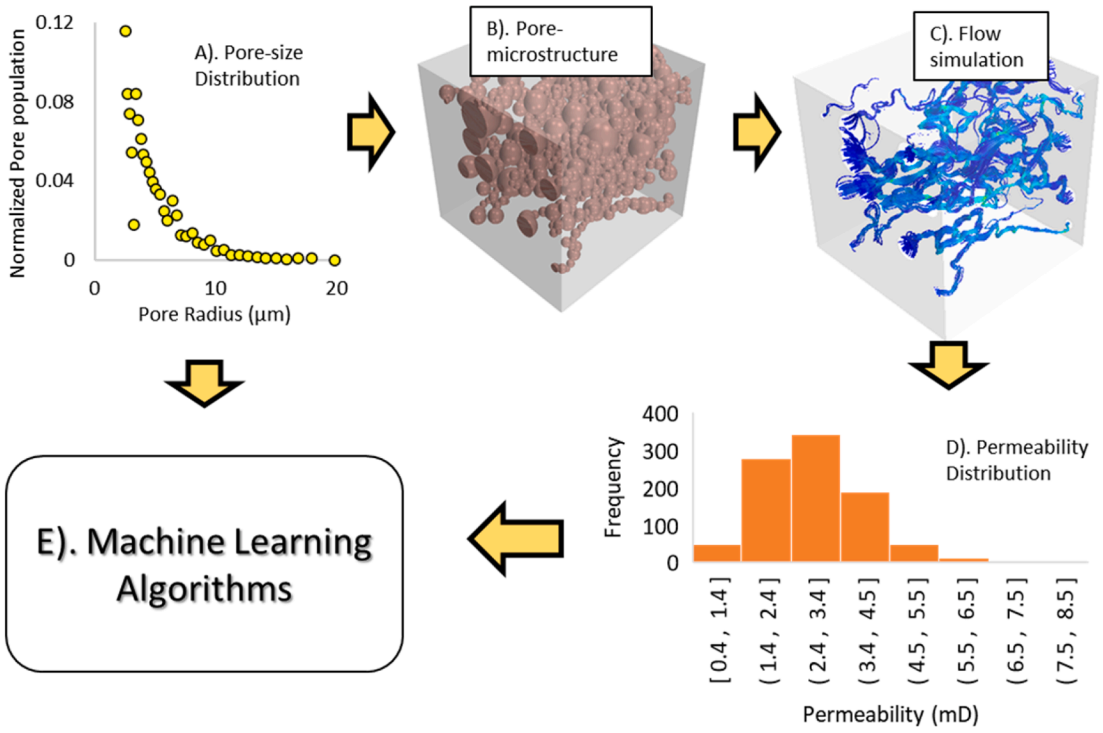


Fig. 2. Workflow of permeability prediction.

rocks can vary significantly in space even within the same outcrop due to diagenesis [29], leading to poor estimates of permeability.

Numerical approaches have been proposed and used to estimate permeability of rocks [30,31]. The main advantage in the use of numerical approaches is the fact that it is not biased since it is dependent on fundamental governing equations of fluid flow [32,33]. However, numerical approaches are dependent on the fidelity of the 3D microstructure image used to conducting flow simulation as well as its representative nature. Different approaches have been employed to construct 3D pore microstructures of rocks from either 2D SEM images or 3D X-ray micro-CT scan images [30,34]. The construction of 3D pore microstructures from 2D SEM images has been done by packing together irregular grains using statistical techniques (e.g., two-point, multiple-point, and simulated annealing) [35,36]. That said, 2D representations of pore microstructural features can vary significantly from their true 3D nature in heterogeneous rocks. The use of FIB-SEM image overcomes this challenge but is a very expensive and time-consuming procedure. Notably, FIB-SEM is extremely small and does not provide a representative distribution of pore microstructural features of rock samples [30,34]. The use of 3D pore microstructures from 3D X-ray micro-CT solves this problem. Though, this technology has a typical resolution of about 1–50 μm which fails to capture smaller pores in rocks as well as pore connectivity [34]. For permeability calculations, numerical methods such as Navier Stokes [33] and Lattice Boltzmann methods [32] provide good estimates of permeability, however, it requires a lot of computational resources which makes its adoption less attractive. These encourages the use of pore network modeling which is less accurate in comparison [37]. This is due to over simplification of pore microstructures making permeability calculations faster and less computationally intensive [37].

To reduce the time and cost of obtaining representative 3D pore microstructure images and conducting respective pore-scale flow simulations, we use machine learning (ML) models instead of pore-scale flow simulations to predict permeability. ML models in the approach is trained with permeability values obtained by a novel stochastic pore-scale simulation approach.

The stochastic pore-scale simulation approach is detailed in our prior paper [16]. This approach entails generating 3D pore microstructures of the same porosity and pore size distribution (PSD) but stochastic pore connectivity. Direct pore-scale simulations of permeability on the stochastically created 3D pore microstructures is used in obtaining the permeability distribution which is analyzed to determine the most probable permeability of a sample with a given porosity and PSD. Direct pore-scale simulations is a common approach to infer porous media properties using pore microstructure images [38,39]. In some cases, it might require thousands of stochastically generated 3D pore microstructures to reach representative elementary values. For practicality, ML is used to reduce the computational cost of direct pore-scale simulations of permeability [37]. The use of ML in geoscience is routine and has been used in making permeability predictions at various scales [40–43]. The features employed in these studies include salts mass concentrations, porosity, lithology, depth, density, grain size, sample color, sample images, formation factor, pore throat diameter, tortuosity, and specific surface area. Algorithms used included linear regression, decision tree, random forest, gradient boosting, support vector machines, and neural network.

Different from previous approaches, our approach relies solely on the fundamental pore microstructural parameters as features while the target (permeability) is physics derived through direct pore-scale simulations through the Navier Stokes equation. A key advantage of our approach over existing ML approaches for pore-scale permeability predictions lies in the use of PSD of rocks instead of pore-scale images which existing models generally use as feature data [37,41,44,45]. This implies that the computational resources required by our ML approach is significantly lower since our feature data is a CSV file and not an image data that could be up to tens of Gigabytes in size which requires large computer RAM and GPUs to run. Furthermore, the use of MICP data as in this study provides continuum scale pore size distribution in rocks at higher resolution compared to direct imaging techniques [30,34,46].

The workflow introduced in this study is easy to implement given the ease of obtaining continuum scale porosity and PSD of rocks from nuclear magnetic resonance (NMR) and mercury injection capillary pressure (MICP) data.

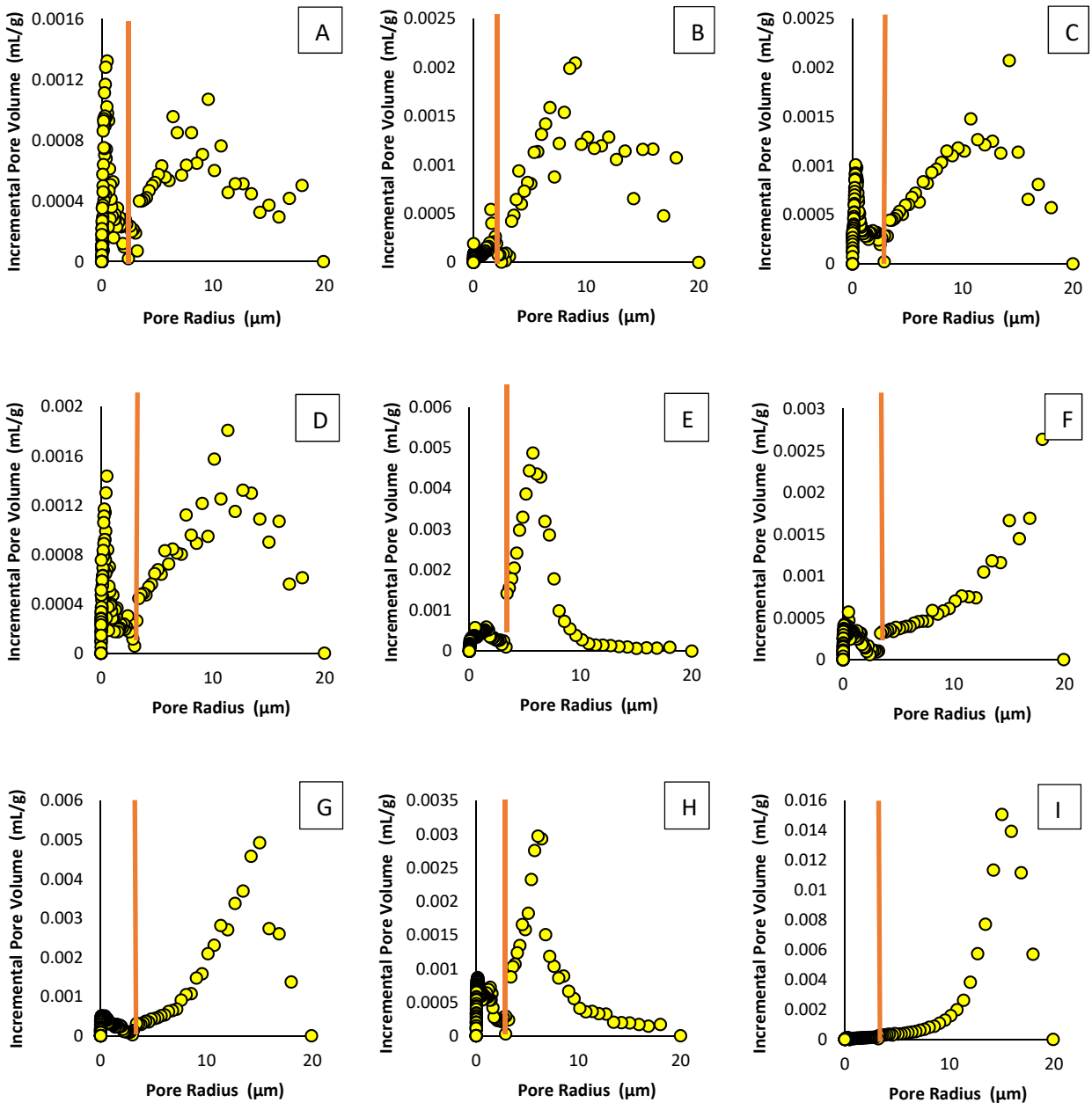


Fig. 3. MICP data of pore radius against incremental pore volume for all core plugs used in this study (Fig. 1).

2. Material and methods

2.1. Workflow

To test the proposed ML approach to predict permeability, nine core plugs were obtained. The description of the core plugs is shown in Fig. 1. Fig. 2 shows the general workflow of this study. The data required in implementing the workflow are PSD and porosity of a rock sample (Fig. 2A) which is used in generating 3D pore microstructures of the same PSD and porosity but different stochastic pore connectivity (Fig. 2B). Pore-scale flow simulations are conducted on the generated 3D pore microstructures (Fig. 2C) which provides a distribution of possible permeability values of the analysed core plug (Fig. 2D). The input data (porosity and PSD) and respective distribution of possible permeability values are used to train a ML model (Fig. 2E). It is noteworthy that the ML model could be trained with combined data of several rock samples.

The permeability distribution obtained from pore-scale simulations was averaged to obtain the permeability for the respective core plug which was compared to measured values as wells as permeability obtained from known empirical model equations. The permeability of the core plugs were obtained from standard laboratory procedures while porosity and PSD of the core plugs were obtained from their MICP data.

2.2. Permeability measurement

Permeability measurements was based on Darcy's law[12] which can be expressed as:

$$K = \frac{Q \times \mu \times L}{A \times \Delta P} \tag{3}$$

where Q is the volumetric flowrate through the inlet or outlet in m^3/s, A is the cross-sectional area perpendicular to the flow at the respective

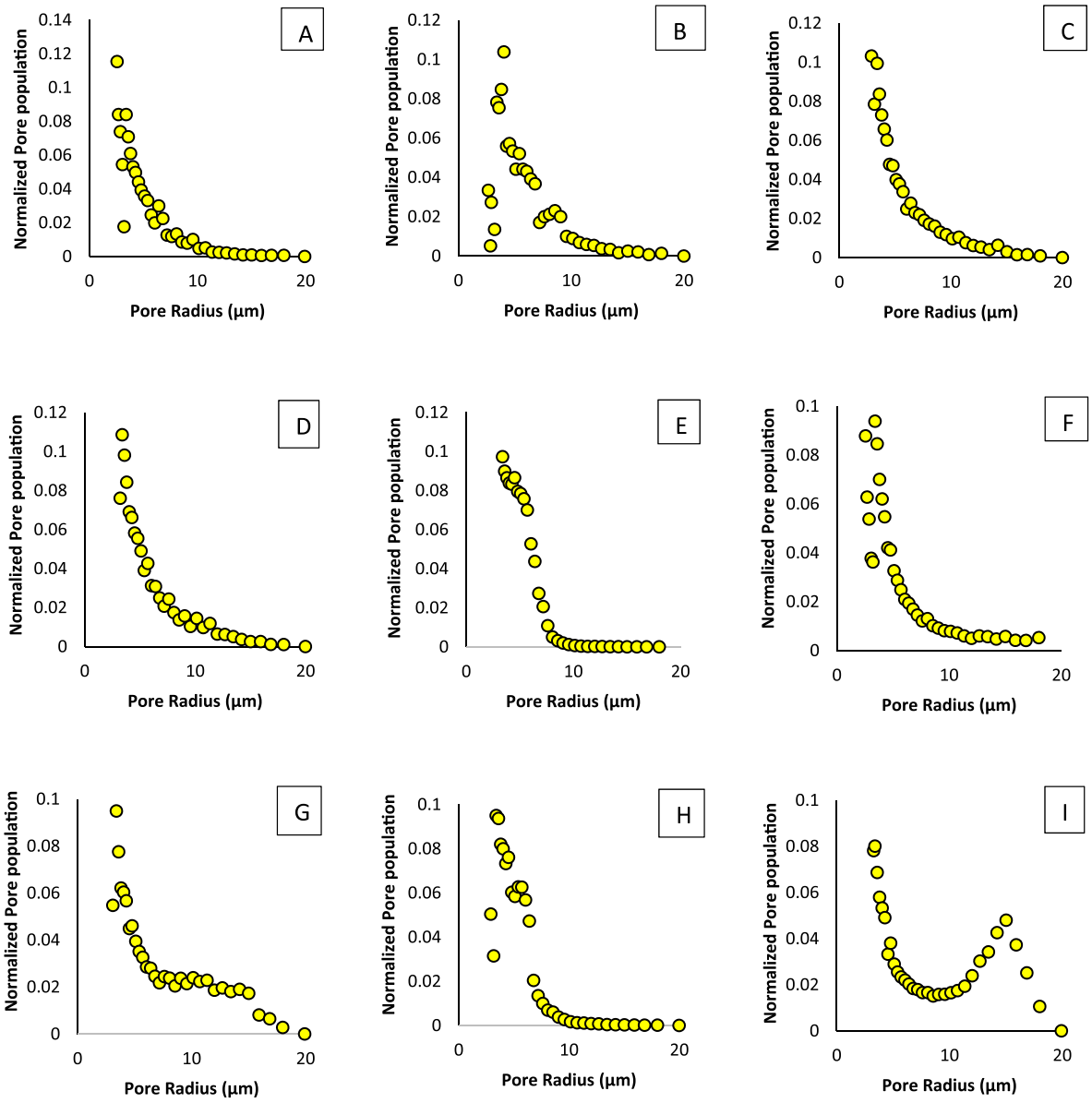


Fig. 4. MICP data plot of pore radius against normalized pore population for all core plugs used in this study (see Fig. 1).

boundary in m^2, K is the absolute permeability of the rock domain in m^2, ΔP is the pressure drop in the principal direction of the fluid in Pa, μ is the dynamic viscosity of the fluid in Pa-s, and L is the length along the direction of the pressure drop in meters. Calculated permeability was converted from m^2 to mD.

For pressure drop measurement, we used a Hassler Type core-holder (RCH-series of Core Laboratory) where water was injected into the core plugs through a 260 dual syringe pump (Teledyne ISCO). Flow rates varied from 0.025 to 0.5 mL/min. Confining pressure was applied to close the space between the core plug and the interior of the core holder to prevent flow of injected fluid around the core plug. Low flow rates were used to avoid flow channelling through the core plug and to ensure that the Reynolds number is very low (less than 1×10^{-4}); enough for the application of Darcy's law to be valid for permeability calculation. A pressure transducer was used to measure the pressure at the inlet while the outlet pressure is known to be atmospheric pressure. The pressure drop across the core plug was observed through time to detect when the system reaches a steady state condition. The experiment is deemed to reach steady state when the pressure drop becomes constant. L and A were measure on the respective core plugs, while the dynamic viscosity of the fluid (water) is known to be 8.90×10^{-4} Pa·s.

2.3. Pore size distribution (PSD) data

PSD data of the core plugs was obtained from MICP data (Fig. 3). MICP is routinely used to deduce PSD in rocks using the Washburn equation [47–53]. To partially account for shadowing or ink-bottle phenomenon in MICP data [34,54], a cut off is applied to the MICP distribution. The cut off is applied at the minimum incremental pore volume (MIPV). The MIPV (Fig. 3) is coined in this study as pore size where the incremental porosity abruptly drops before attaining 100 % mercury saturation. A cut off is applied at MIPV because the pressure to get to the pores beyond this point is already very high; hence the entire pore system unsaturated at this point is erroneously captured as tiny pores [34,54]. Furthermore, the low magnitude of incremental volume at the MIPV can be interpreted as a semi-isolated pore system which

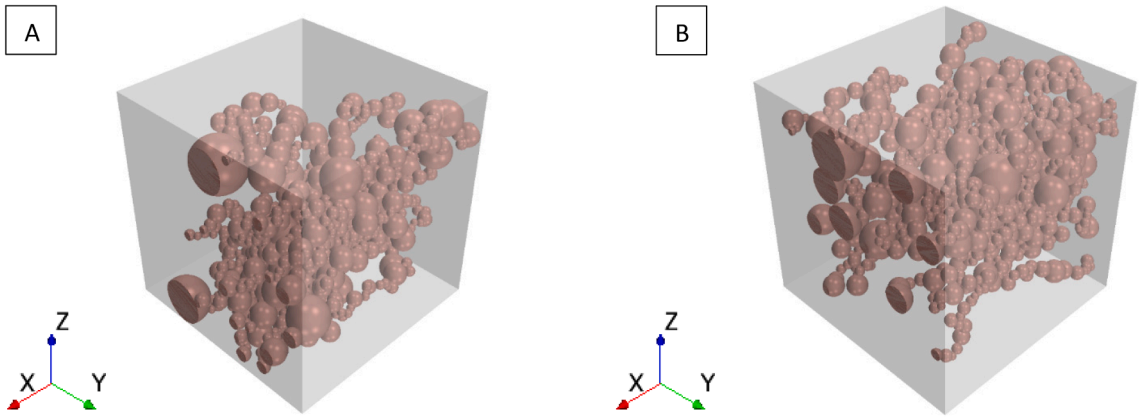


Fig. 5. Representative stochastically generated 3D pore microstructures showing different possible pore connectivity (brown) scenarios in core plug A (Fig. 1).

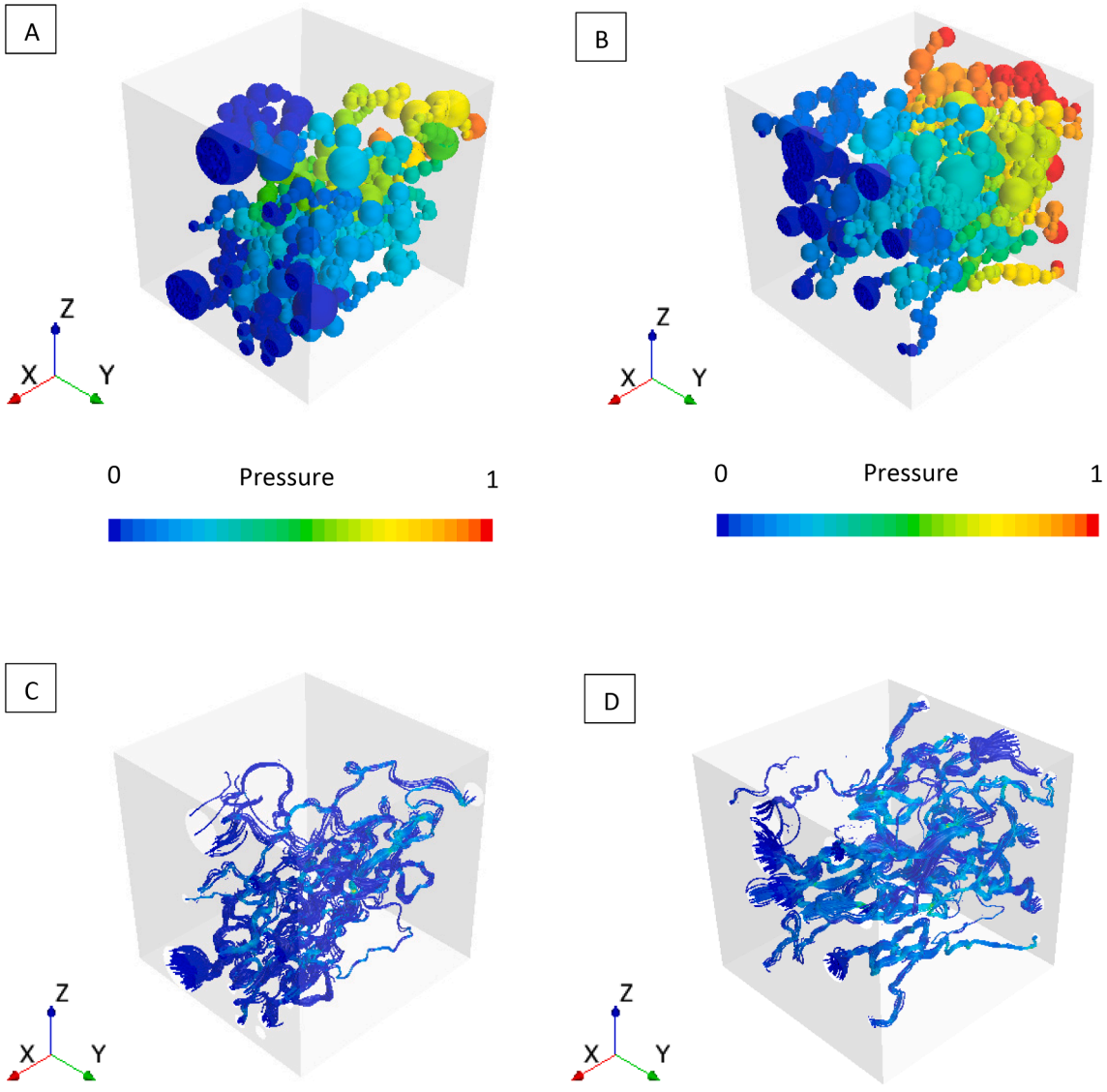


Fig. 6. Representative pressure distribution (A and B) and corresponding streamlines (C and D) through stochastically generated pore microstructures from PSD data (Fig. 4.A and 4.B). Flow is in the positive-X direction.

might contribute to the overall flow through the core plug. In absence of accurate data for this region, we assume that the region has the same pore size distribution as the remainder of the core plug. It is noteworthy, that the applied cut-off also discards tiny pores in the core plugs.

In this study, the pore geometry is simplified to a sphere. Data provided from MICP analysis of core plugs is a plot of pore radius against

incremental pore volume (Fig. 3). To obtain the pore population of the core plugs, the incremental pore volume is divided by the corresponding volume of a single pore (Eq. (4)). Pore population was normalized (Eq. (5)) to ease comparison among the core plugs (Fig. 5).

$$P = \frac{P_v}{\frac{4}{3}\pi r^3} \tag{4}$$

$$P_p = \frac{P}{\sum P} \tag{5}$$

where P_v is the incremental pore volume from MICP data, r is the corresponding pore radius, $\sum P$ is the total number of pores in the analysed core plug, P is the number of pores of certain pore radius, and P_p is the normalized pore size distribution.

Fig. 4 shows the PSD of all the core plugs to have lognormal distribution. Our core plugs include unimodal (Core plug A-H) and bimodal (Core plug I) PSD to test the versatility of the workflow. The MICP analysis also provided porosity ($\varnothing_m$). It's important to note the porosity provided by MICP is effective porosity given that all the pores must be accessible by injected mercury.

2.4. Stochastic generation of 3D pore microstructures

The stochastic generation of 3D pore microstructures (Fig. 5) follows the workflow employed by Ishola et al., 2022 [16] where pore size distribution, effective porosity, and a spherical pore geometry were used to create multiple realizations of 3D pore microstructures of the same effective porosity, PSD, and pore shape while stochastically connecting the pores. The fixed effective porosity, PSD, pore frequency, and pore shape helps consider the effect of a wide range of possible pore connectivity scenario in each core plug [16]. This is vital because pore connectivity cannot be accurately deduced in a representative core plug despite been a key control on permeability [16].

For each core plug, the number of stochastically generated 3D pore microstructures varied from 845 to 933. In total, 8,123 pore microstructures were generated across the nine core plugs. Summary data of the pore microstructures can be found in the supplementary information (Table 1S and 2S). Full details on the stochastic generation of 3D pore microstructures and precautions taken to ensure the representative nature of the pore microstructural parameters is detailed in our preceding publication [16].

2.5. Pore-scale simulations

Pore-scale simulations of flow (Fig. 6) through the stochastically generated 3D pore microstructures is executed with STAR-CCM+® computational fluid dynamics software, using its finite volume methodology to solve the mass continuity equation [16]:

$$\vec{\nabla}\cdot\rho\vec{V} + \frac{\partial\rho}{\partial t} = 0 \tag{6}$$

and Navier-Stokes momentum equation:

$$\frac{\partial\vec{V}}{\partial t} + \left(\vec{V}\cdot\vec{\nabla}\right)\vec{V} = -\frac{1}{\rho}\vec{\nabla}P + \nu\vec{\nabla}^2\vec{V} \tag{7}$$

where ρ is density of the fluid, ν is kinematic viscosity, P is pressure, t is time, and $\vec{V}$ is fluid velocity.

This software has been successfully used for flow simulations in similar 3D pore microstructures and has been proven effective in simulating pore-scale flow and solute transport process [55–57]. The implementation of pore-scale simulation in this study also follows the workflow used by Ishola et al., 2022 [16] with the only difference being that the minimum cell size of the mesh was lowered to 0.375 μm from 0.75 μm to provide higher level of detail in the flow domain. The average number of cells across generated 3D pore microstructures is 1 million and the computational time per simulation is about 11 min. We employed Oklahoma State University's Pete supercomputing facilities (32 cores and 96 GB of RAM for each run) to execute the numerical portion of this study.

2.6. Permeability calculation

Numerical simulation of flow through each 3D pore microstructure at a Reynold number less than 0.0001 makes permeability estimation possible through Darcy's equation (Eq. (3)) [7,12,31] and is rewritten as [16]:

$$K = \frac{M \times \mu \times L}{\Delta P \times A \times \rho} \tag{8}$$

where K is the absolute permeability of the rock domain in m^2, ΔP is the pressure drop in the principal direction of the fluid in Pa, μ is the dynamic viscosity of the fluid in Pa-s, M is the mass flowrate through the inlet or outlet in kg/s, ρ is the density of the fluid flowing through the medium in kg/m^3, A is the cross-sectional area perpendicular to the flow at the respective boundary in m^2, and L is the length along the direction of the pressure drop in meters. Since properties of the fluid (density and viscosity of water) and pressure drop are constant, permeability is calculated from changes in mass flow rate and length scale (L^{-1}) of the 3D pore microstructure[16]. Calculated permeability was converted from m^2 to mD.

2.7. Machine learning implementation

Several ML algorithms were trained with pore microstructural parameters obtained from the generated 3D pore microstructures as features and corresponding permeability estimated from pore-scale simulations as the target. For ML implementation, the entire data set—permeability estimated from pore-scale flow simulations of all 3D pore micro structures generated for the nine core plugs—was randomized, and split into training and test set in ratio 7:3. The pore microstructural parameters (features) used for ML implementation are the average pore-throat size, standard deviation of pore-throat size distribution (PTSD), 1st percentile of PTSD, 5th percentile of PTSD, 10th percentile of PTSD, 25th percentile of PTSD, 50th percentile of PTSD, 75th percentile of PTSD, 90th percentile of PSD, 99th percentile of PTSD, porosity, domain size of the pore microstructure, number of pores connected to the inlet (conin) and outlet (conout) flow faces of the pore microstructure, and minimum of conin and conout of the respective pore microstructures. The last three features are introduced in this study, and they are deemed to capture pore connectively as they quantify the true number of alternative paths for a particle to from inlet to outlet. The target in this study is permeability. The ML algorithms considered in this study are Linear Regression, Random Forest, Support Vector Machine, Gradient Boosting, and Artificial Neural Network [58]. The algorithm with the best performance in our test is the Gradient Boosting (GB) algorithm based on mean absolute percentage error [59] between permeability obtained from the direct pore-scale simulations and permeability obtained from respective ML algorithms. Hence, only the parameters related to GB modelling are mentioned in this paper. A summary on the implementation of the remaining five ML algorithms and their relative performance is in the supplementary information (Fig. 1S and Table 3S). In the pre-processing stage of the data for machine learning deployment, the data was standardized to account for the scale difference in the features [60,61]. In this study, the data was standardized by:

$$s = \frac{x - \bar{x}}{\sigma} \tag{9}$$

where s is the value after standardization, x is the original value of feature, $\bar{x}$ is the feature average, and σ is the standard deviation of the

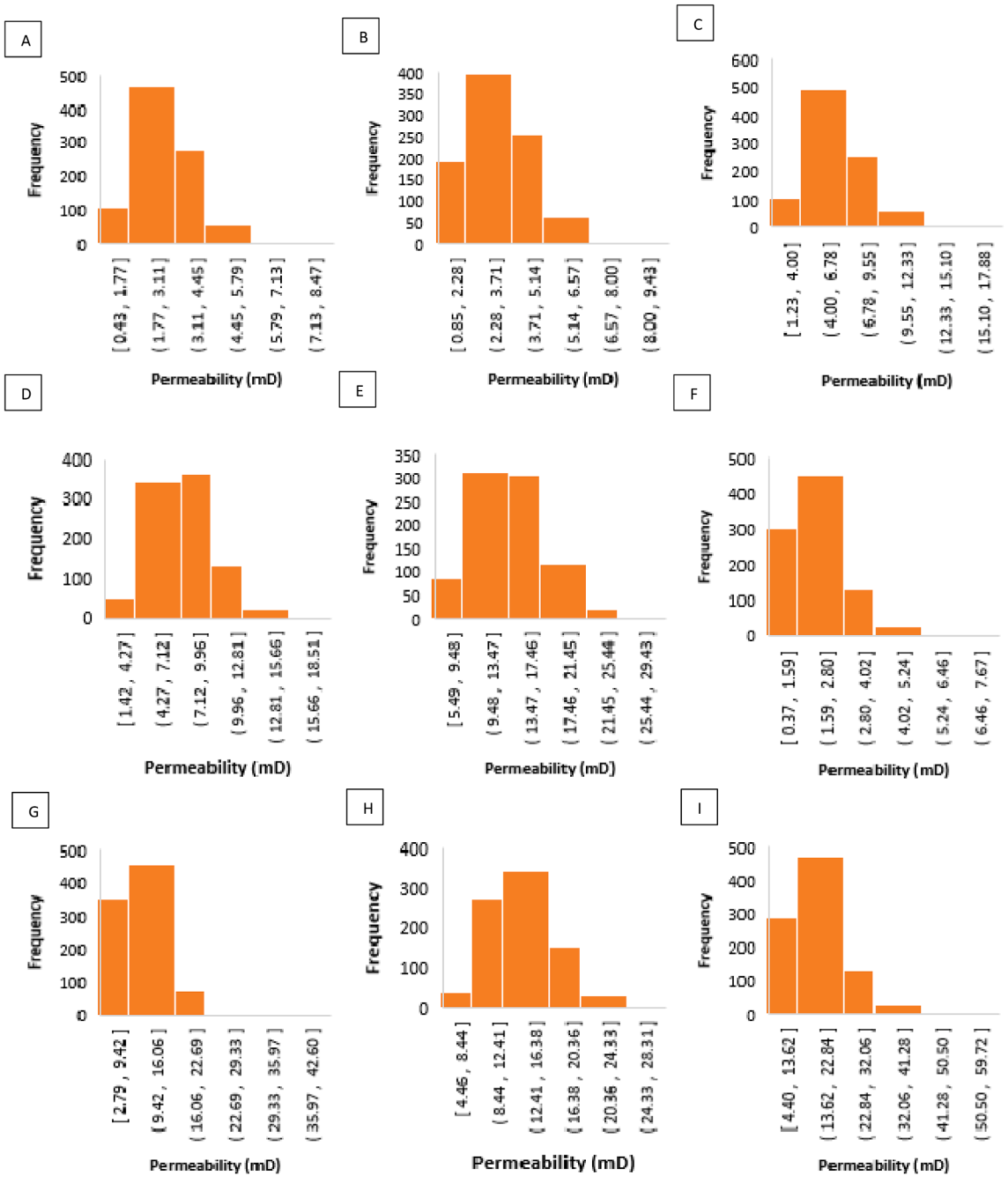


Fig. 7. Permeability distribution from stochastic pore scale simulations through the generated 3D pore-microstructures for each core plug.

Table 1
Comparison of permeability (mD) between previous deterministic approaches and the stochastic pore-scale simulation approach used in this study with mean absolute percentage error (MAPE).

Sample	A	B	C	D	E	F	G	H	I	MAPE	Rank
Experimental permeability	20.96	3.19	15.13	35.31	8.73	17.03	15.13	6.69	52.46	NA	NA
Winland[24]	9.29	68.75	46.16	40.51	68.66	53.14	240.16	36.18	536.31	676%	4th
Swanson[27]	12.68	55.93	42.19	40.02	67.83	44.88	232.4	26.19	1593.46	821%	5th
Wells-Amaefule[28]	2.28	7.24	5.82	5.58	8.41	6.10	21.94	4.01	98.10	67%	2nd
Kamath[23]	24.29	79.41	63.41	60.8	92.62	66.62	247.46	43.34	1150.05	914%	6th
Dastidar[22]	5.22	36.49	12.24	12.14	57.81	6.05	49.02	11.44	1640.37	573%	3rd
Stochastic simulation approach	2.87	3.33	6.21	7.75	14.00	2.05	10.94	14.00	17.60	64%	1st

Table 2
The respective absolute error (mD) across the nine core plugs in Table 1.

Sample	A	B	C	D	E	F	G	H	I
Winland[24]	11.67	65.56	31.03	5.2	59.93	36.11	225.03	29.49	483.85
Swanson[27]	8.28	52.74	27.06	4.71	59.10	27.85	217.27	19.5.0	1541.00
Wells-Amaefule[28]	18.68	4.05	9.31	29.73	0.32	10.93	6.81	2.68	45.64
Kamath[23]	3.33	76.22	48.28	25.49	83.89	49.59	232.33	36.65	1097.59
Dastidar[22]	15.74	33.30	2.89	23.17	49.08	10.98	33.89	4.75	1587.91
Stochastic simulation approach	18.09	0.14	8.92	27.56	5.27	14.98	4.19	7.31	34.86

Table 3
Data summary of the empirical deterministic permeability models explored in this study (modified after Comisky et al. [65]).

Method	Year	Number of Samples	Sample Source	Permeability Measurement Type (s)	Permeability Range (mD)	Data Density*
Winland[24]	1980	322	Mixed	Air, Klinkenberg corrected	N/A	N/A
Swanson[27]	1981	56	Mixed	Brine	0.002–1000	0.06
Wells-Amaefule[28]	1985	35	Siliciclastic	pulse decay	0.00002–70	0.50
Kamath[23]	1992	301	Mixed	Klinkenberg corrected	1–2000	0.10
Dastidar[22]	2007	150	Siliciclastic	Klinkenberg corrected	0.0001–10000	0.02

* Estimated by diving the number of samples by the permeability range.

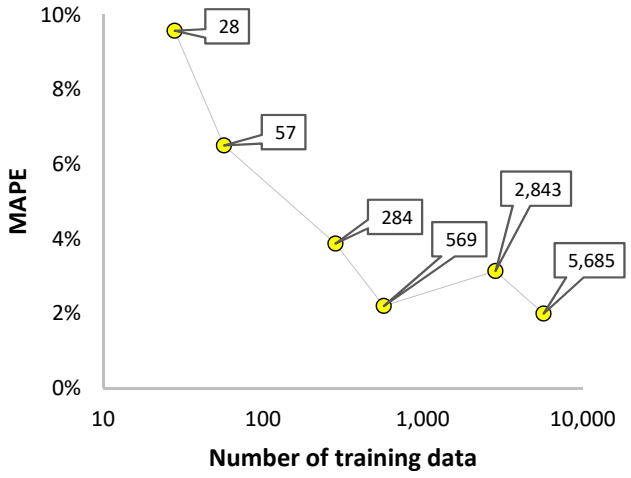


Fig. 8. MAPE of permeability of all nine core plugs used in this study. MAPE by core plug is shown in the supplementary information (Fig. 2S). The data highlighted in plot is number of training data.

features. Hence, all the features have a mean of zero and a standard deviation of one.

GB was implemented on Scikit-learn [61]. The maximum depth in the GB algorithm was varied between 3 and 10 while the number of estimators was varied from 10 to 1000. The learning rate used in this study ranged from 0.0001 to 1 with other parameters in the GB algorithm kept at default. GridSearchCV [61] was used to select the best combination of maximum depth and number of estimators. The GB model was trained with data from all nine core plugs and combination of hyperparameters was selected based on a five-fold cross validation. In the second phase of GB deployment, the number of training data was systematically reduced from 100% to 0.5% to test the minimum number of training data required to make relatively good permeability predictions. It's worth highlighting that the test data remained constant for objective comparison.

2.8. Statistical comparison of results

The permeability of each core plug was estimated by averaging the permeability of the 3D pore microstructures that were stochastically generated from the PSD and porosity data (Fig. 4) [62–64]. The permeability of each core plug estimated using this method was compared against permeability estimated using five notable model equations through the mean absolute percentage error (MAPE) [59]. MAPE is the average of the absolute percentage error between measured permeability and estimated permeability for n pore plugs. It can be expressed as:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{k_{m,i} - k_{e,i}}{k_{m,i}}\right| \times 100 \tag{10}$$

where $k_{m,i}$, and $k_{e,i}$ is the measured permeability and estimated permeability of core plug i, while n is the total number of core plugs. MAPE was used in this study because it takes magnitude of the measured permeability into account, hence, an objective comparison across different scales can be made.

2.9. Previous approaches to permeability estimation

For comparison with the approach used in this study, permeability was estimated for all nine core plugs using five notable model equations that employ MICP data. This include Winland [24], Swanson [27], Wells-Amaefule [28], Kamath [23], and Dastidar models [22]. The respective equations are given by:

$$k_{Winland} = 49.4 * R_{35}^{1.7} * \varnothing^{1.47} \tag{11}$$

$$k_{\text{Swanson - brine}} = 355 * \left[\frac{s_b}{P_C}\right]_A^{2.005} \tag{12}$$

$$k_{\text{Wells - Amaefule}} = 30.5 * \left[\frac{s_b}{P_C}\right]_A^{1.56} \tag{13}$$

$$k_{\text{Kamath}} = 347 * \left[\frac{s_b}{P_C}\right]_A^{1.60} \tag{14}$$

$$k_{Dastridar} = 4073 * R_{wgm}^{1.64} * \varnothing^{3.06} \tag{15}$$

where s_b is the percent bulk volume occupied by mercury, P_C is the mercury capillary pressure (Psia), A is the maximum amplitude, R_{35} is 35% mercury saturation of pore volume, $\varnothing$ is porosity (fraction), and R_{wgm} is the geometric mean of pore sizes. Eqs. (11)-(14) are dependent on the empirical correlation between a critical pore throat sizes where mercury infiltration is deemed to connect the entire pore system of rock sample while Eq. (15) considered the size of all the pore throat sizes that

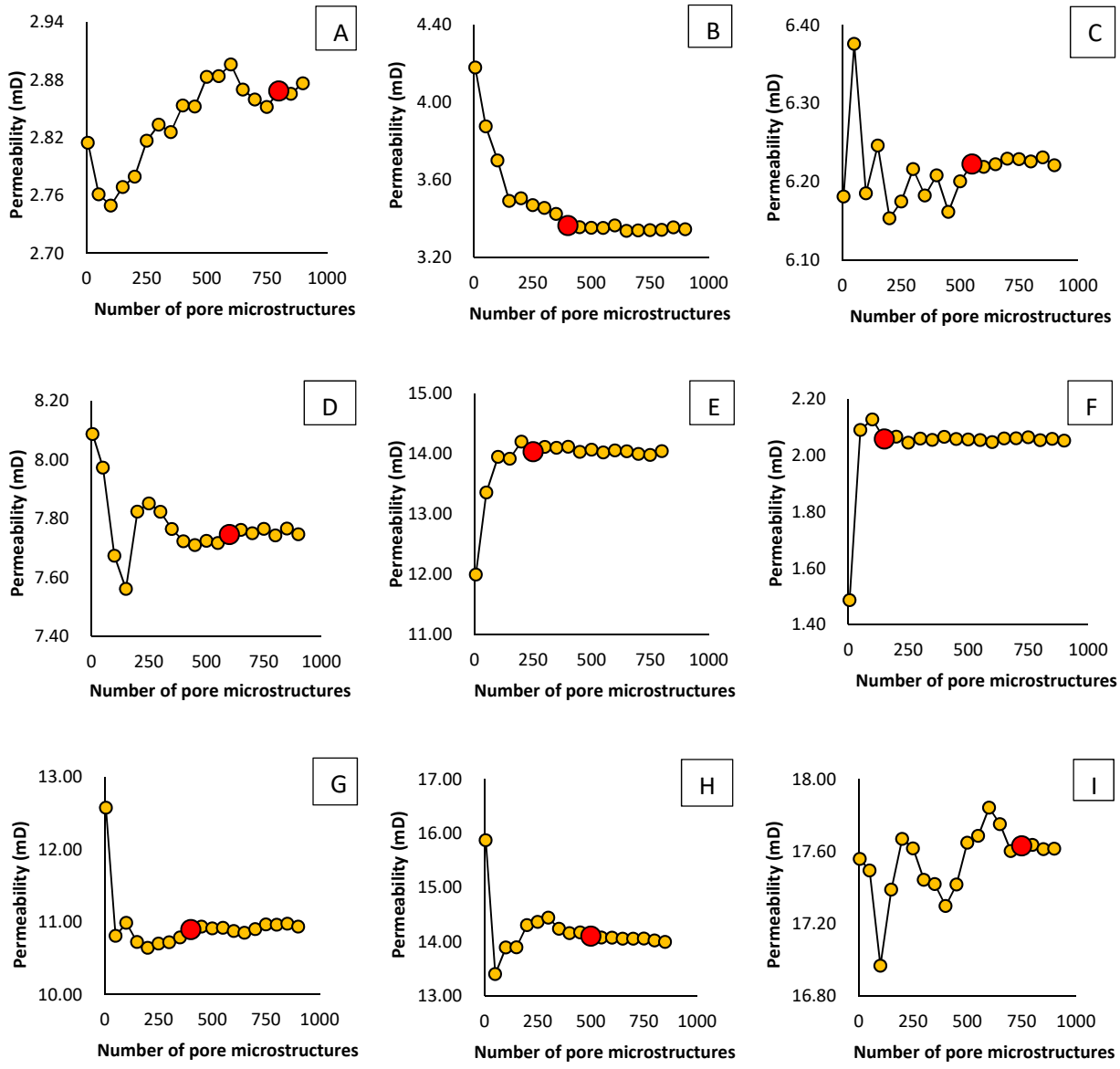


Fig. 9. Change in calculated permeability with number of 3D pore microstructures used for pore-scale flow simulations. The red point indicates the representative permeability value of the respective core plug. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

are in a rock sample by taking their geometric mean.

3. Results and discussion

3.1. Stochastic pore-scale simulations of permeability

Direct pore-scale simulations of permeability for the 3D pore microstructures stochastically generated from the PSD curves (Fig. 4) yield a distribution of possible permeability values for the respective core plug (Fig. 7). The permeability for each core plug is calculated via arithmetic mean of possible permeability values. In comparison to existing deterministic model equations attempted in this study (Tables 1 and 2), computed permeability for the core plugs is generally closer to measured permeability of the respective core plug. Table 1 shows the relative rank of all the methods based on MAPE for the nine core plugs used in this study. The Wells-Amaefule model [28] ranks second place after our approach and is the only other approach with an MAPE less than 100%. Wells-Amaefule model is only an expansion of the empirical correlation introduced by Swanson [27] to include tight rocks with permeability as low as 0.00002 mD [21]. That said, the maximum permeability in the empirical correlation by Wells-Amaefule is 70 mD which is within a reasonable range compared to the permeability data used in our study (Fig. 1 and Table 3). It also implies that the Wells-Amaefule model might not work well with higher permeability data since the empirical model is fitted to a narrow permeability data range. That said, the permeability data density is relatively high compared to the other empirical model equations implemented in this study (Table 3). it is important to note that the permeability data used in the Wells-Amaefule model is gas permeability. However, the authors argued that it is on the same order of magnitude as Klinkenberg corrected and liquid permeability given the conditions under which the values are obtained [28]. In third place of MAPE ranking is the Dastidar model [22]. The Dastidar model [22] is fundamentally different from the other empirical model equations considered in this study. As show in Eq. (15), it takes the entire pore size

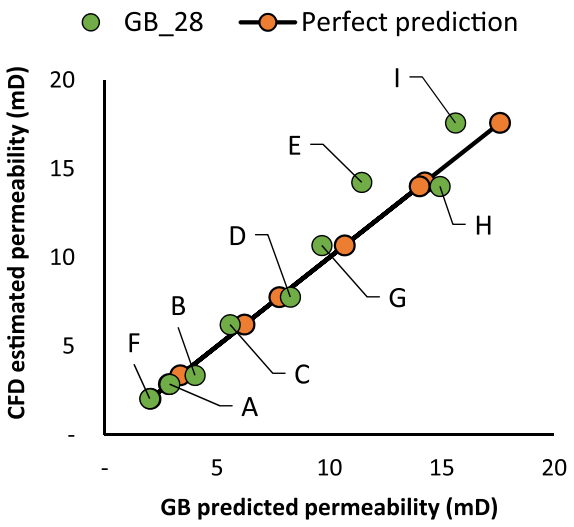


Fig. 10. A comparison between permeability predictions made using a GB model trained with 28 pore microstructural parameters (green) and an hypothetical perfect prediction (orange). The higlighted numbers are respective core plug IDs. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

distribution into account by using their geometric mean. In contrast, the other four model equations (Eqs. 11–14) use a threshold in mercury saturation through the rock samples. However, the MAPE of permeability obtained using the Dastidar model was 573%, which is relatively high. This could be due to three reasons: 1) The Dastidar model gives more weight to porosity than the average pore-throat size (Eq. (15)). Hence, porosity is the controlling parameter in this model. Porosity is known to have a relatively poor relationship with permeability [18]. 2) The weights used in Dastidar model (Eq. (15)) is derived from incremental porosity [21,22] and not the pore populations. We believe pore population better reflects true frequency of pores in a rock sample. 3) The permeability data density used in developing the Dastidar model is relatively low (Table 3). The Winland approach [24] ranks 4th place in this study with an MAPE of 676%. The performance of this model could be due to the inclusion of uncorrected air permeability in modelling which is known to overestimate permeability [21]. 240 of the 322 permeability data points used in developing the Winland model are uncorrected air permeability (Table 3). The Swanson-brine model [27] ranked 5 out of the 6 methods in this study. The Swanson- brine model [27] was established with permeability ranging from 0.02 mD to 1000 mD, however, the permeability data density is low (Table 3). Given that the main difference between Wells-Amaefule model[28] and the Swanson- brine model[27] is a significant reduction in permeability data density in Swanson- brine model [27] (Table 3), we can conclude that this is responsible for the drop off in MAPE from 67% in Wells-Amaefule model [28] to 821% in Swanson-brine model [27]. It is noteworthy that permeability range used to establish the Wells-Amaefule model equation [28] (Table 3) is concentrated within the permeability data range of this study (Fig. 1) which might also influence the performance of Wells-Amaefule model [28]. This underlines a pitfall in the use of empirical model equations because it is usually suited to the data distribution from which it was derived and can be unreliable when the permeability data density is poor. Kamath model [23] ranked 6 out of 6 in in approaches tested in the study with a MAPE of 914%. Considering that permeability prediction is consistently overestimated across all core plugs (Table 1), it is plausible that the performance is driven by the datasets used in the study. Based on the relative performance of the model equations used in this study, we can conclude that, empirical permeability model equations are adversely impacted by permeability data density, the use of uncorrected air permeability, reliance on porosity as the key parameter, and the distribution of permeability data used in fitting the relationship. The use of the stochastic pore-scale simulation approach [16] overcomes these limitations because it mimics mechanistic processes that occur in rocks through the governing equation of fluid flow, and it considers different pore connectivity scenarios that can result in the same PSD and porosity.

3.2. Computational reduction with machine learning

Fig. 8 shows only a 10% difference between the permeability values computed from pore-scale flow simulations and a GB model trained with pore microstructural parameters (features) of 28 pore microstructures. It is worth noting that the total number of pore microstructures required to obtain a representative permeability value from pore-scale flow simulations across the nine core plugs employed in this study is ~4400 (Fig. 9). This implies that the time required to obtain representative permeability prediction across the core plugs is reduced from 34 days (4,400 3D pore microstructures) to 6 h (28 3D pore microstructures). If a pretrained GB model is available, permeability predictions will take less than a second on a 16 GB RAM computer. Fig. 8 also shows that permeability predictions do improve with the size of training permeability data. Therefore, there is need for trade-off between available computing resources and accuracy the GB model. That said, Fig. 10 shows 28 pore microstructures to be sufficient (green data points) as it is close to a perfect prediction (orange data points). However, if the objective is to replicate representative permeability values computed from pore-scale flow simulations with near perfection, a larger number of training permeability data is beneficial. It is worth nothing that the only information needed to implement this ML approach is porosity and PSD which can be obtained from X-ray imaging methods, MICP, and NMR data.

4. Conclusions

In this study we have employed a stochastic pore-scale simulation approach to predict the permeability of nine core plugs of both carbonate and siliciclastic sources. Our stochastic pore-scale simulation approach is more accurate in predicting permeability compared to five notable deterministic model equations attempted in this study. To reduce the computational cost of predicting permeability by our stochastic pore-scale simulation approach, we used Gradient Boosting algorithm. This ML approach reduces the number of simulations needed to obtain representative permeability of all nine core plugs from 4,400 to 28, significantly reducing required computational resources and time. There are at least three key applications for our results: 1) a pretrained ML model using our approach can be incorporated with NMR information obtained while drilling to provide permeability information of penetrated formations, instantaneously; 2) a pretrained ML model using our approach could provide a simple tool to make quick assessments of rock samples before full laboratory experiments are applied to get true permeability of core plugs; and 3) the approach used in this study can help improve the performance and accuracy of fluid flow and reactive transport simulation computer programs. In general, the result of this study is beneficial in resource exploration as well as environmental protection and management.

CRediT authorship contribution statement

Olubukola Ishola: Methodology, Investigation, Formal analysis, Software, Writing – original draft. **Javier Vilcáez:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Declaration of Competing Interest

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

Acknowledgments

We appreciate the Association of Petroleum Geologist Foundation (AAPG), Oklahoma Geological Foundation (OGF), and the Oklahoma City Section of Society of Petroleum Engineers (SPE) for partial financial support of the Ph.D. research of the first author. This is Oklahoma State University Boone Pickens School of Geology contribution number 2022-130. Some of the computing for this project was performed at the High-Performance Computing Center at Oklahoma State University (OSU).

Funding

This work was supported by the National Science Foundation under Grant HS-2041648.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fuel.2022.124044.

References

[1] Beisman JJ, Maxwell RM, Navarre-Sitchler AK, Steefel CI, Molins S. ParCrunchFlow: an efficient, parallel reactive transport simulation tool for physically and chemically heterogeneous saturated subsurface environments. Comput Geosci 2015;19(2):403–22.
[2] Harbaugh AW. MODFLOW-2005: The U.S. Geological Survey modular ground-water model–the ground-water flow process. *Techniques and Methods*. – ed. 2005.
[3] Xu T, Sonnenthal E, Spycher N, Pruess K. TOUGHREACT—A simulation program for non-isothermal multiphase reactive geochemical transport in variably saturated geologic media: Applications to geothermal injectivity and CO2 geological sequestration. Comput Geosci 2006;32(2):145–65.
[4] Friedman M. Porosity, permeability, and rock mechanics-a review. The 17th US Symposium on Rock Mechanics. 1976.
[5] Zhang L. Aspects of rock permeability. Frontiers of Structural and Civil Engineering 2013;7(2):102–16.
[6] Algive L, Békri S, Nader FH, Lerat O, Vizika O. Impact of Diagenetic Alterations on the Petrophysical and Multiphase Flow Properties of Carbonate Rocks Using a Reactive Pore Network Modeling Approach. Oil & Gas Science and Technology – Revue d'IFP Energies nouvelles 2012;67(1):147–60.
[7] Chen Y, Shen C, Lu P, Huang Y. Role of pore structure on liquid flow behaviors in porous media characterized by fractal geometry. Chem Eng Process Process Intensif 2014.
[8] Cheng Z, Ning Z, Wang Q, Zeng Y, Qi R, Huang L, et al. The effect of pore structure on non-Darcy flow in porous media using the lattice Boltzmann method. J Petrol Sci Eng 2019;172:391–400.
[9] Mohebbi A, Kaydani H. Permeability Estimation in Petroleum Reservoir by Meta-heuristics: An Overview. In: Artificial Intelligent Approaches in Petroleum Geosciences. Cham: Springer International Publishing; 2015. p. 269–85.
[10] Kozeny J. Ueber kapillare Leitung des Wassers im Boden. Sitzungsber Akad Wiss 1927;136:271–306.
[11] Ozgumus T, Mobedi M, Ozkol U. Determination of Kozeny Constant Based on Porosity and Pore to Throat Size Ratio in Porous Medium with Rectangular Rods. Engineering Applications of Computational Fluid Mechanics 2014;8(2):308–18.
[12] Darcy H. Les fontaines publiques de la Ville de Dijon V Dalmont. 1856.
[13] Sutera SP, Skalak R. The history of Poiseuille's law. Annu Rev Fluid Mech 1993;25(1):1–20.
[14] Ghanbarian B, Hunt AG, Ewing RP, Sahimi M. Tortuosity in Porous Media: A Critical Review. Soil Sci Soc Am J 2013;77(5):1461–77.
[15] Carman P. Fluid flow through granular beds. Chem Eng Res Des 1937;15:150–66.
[16] Ishola O, Alexander A, Vilcáez J. Statistical and neural network analysis of the relationship between the stochastic nature of pore connectivity and flow properties of carbonate rocks (In review).
[17] Sun H, Vega S, Tao G. Analysis of heterogeneity and permeability anisotropy in carbonate rock samples using digital rock physics. J Petrol Sci Eng 2017;156:419–29.
[18] Westphal H, Surholt I, Kiesl C, Thern HF, Kruspe T. NMR Measurements in Carbonate Rocks: Problems and an Approach to a Solution. Pure appl geophys 2005;162(3):549–70.
[19] Timur A. An Investigation Of Permeability, Porosity, & Residual Water Saturation Relationships For Sandstone Reservoirs. SPWLA-1968-vIXn4a2 1968;9(04):10.
[20] Kenyon WE. Nuclear magnetic resonance as a petrophysical measurement. Nucl Geophys 1992;6(2):153–71.
[21] Vilcáez J, York J, Youssef N, Elshahed M. Stimulation of methanogenic crude oil biodegradation in depleted oil reservoirs. Fuel 2018;232:581–90.
[22] Dastidar R, Sondergeld CH, Rai CS. An improved empirical permeability estimator from mercury injection for tight clastic rocks. Petrophysics 2007;48(3):186–90.
[23] Kamath J. Evaluation of Accuracy of Estimating Air Permeability from Mercury Injection Data. SPE Form Eval 1992;7(4):304–10.
[24] Kolodzie S, Jr. Analysis of Pore Throat Size and Use of The Waxman-Smits Equation To Determine Ooip In Spindle Field, Colorado. SPE Annual Technical Conference and Exhibition: Dallas, Texas, Society of Petroleum Engineers 1980:10.
[25] Pittman ED. Relationship of porosity and permeability to various parameters derived from mercury injection-capillary pressure curves for sandstone. AAPG Bull 1992;76(2):191–8.
[26] Purcell W. Capillary Pressures - Their Measurement Using Mercury and the Calculation of Permeability Therefrom. J Petrol Technol 1949;1(2):39–48.
[27] Swanson BF. A Simple Correlation between Permeability and Mercury Capillary Pressures. J Petrol Technol 1981;33(12):2498–504.
[28] Walls JD, Amaefule JO. Capillary Pressure and Permeability Relationships in Tight Gas Sands SPE/DOE. Low Permeability Gas Reservoirs Symposium 1985.
[29] Choquette PW, Pray LC. Geologic nomenclature and classification of porosity in sedimentary carbonates. The American Association of Petroleum Geologists Bulletin 1970;54(2):207–44.
[30] Blunt MJ, Bijeljic B, Dong Hu, Gharbi O, Iglauer S, Mostaghimi P, et al. Pore-scale imaging and modelling. Adv Water Resour 2013;51:197–216.
[31] Vilcáez J, Morad S, Shikazono N. Pore-scale simulation of transport properties of carbonate rocks using FIB-SEM 3D microstructure: Implications for field scale solute transport simulations. J Nat Gas Sci Eng 2017;42:13–22.
[32] Chen S, Doolen GD. Lattice Boltzmann method for fluid flows. Annu Rev Fluid Mech 1998;30(1):329–64.
[33] Chorin AJ. Numerical solution of the Navier-Stokes equations. Math Comp 1968;22(104):745–62.
[34] Xiong Q, Baychev TG, Jivkov AP. Review of pore network modelling of porous media: Experimental characterisations, network constructions and applications to reactive transport. J Contam Hydrol 2016;192:101–17.
[35] Okabe H, Blunt MJ. Prediction of permeability for porous media reconstructed using multiple-point statistics. Phys Rev E Stat Nonlin Soft Matter Phys 2004;70(6 Pt 2):066135.
[36] Manwart C, Torquato S, Hilfer R. Stochastic reconstruction of sandstones. Phys Rev E – Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics 2000;62(1):893–9.
[37] Tembely M, AlSumaiti AM, Alameri W. A deep learning perspective on predicting permeability in porous media from network modeling to direct simulation. Comput Geosci 2020;24(4):1541–56.
[38] Kohanpur AH, Chen Y, Valocchi AJ. Using Direct Numerical Simulation of Pore-Level Events to Improve Pore-Network Models for Prediction of Residual Trapping of CO2. Frontiers in Water. 2022. 3.
[39] Konangi S, Palakurthi NK, Karadimitriou NK, Comer K, Ghia U. Comparison of pore-scale capillary pressure to macroscale capillary pressure using direct numerical simulations of drainage under dynamic and quasi-static conditions. Adv Water Resour 2021;147:103792.
[40] Erofeev A, Orlov D, Ryzhov A, Koroteev D. Prediction of Porosity and Permeability Alteration Based on Machine Learning Algorithms. Transp Porous Media 2019;128(2):677–700.
[41] Graczyk KM, Matyka M. Predicting porosity, permeability, and tortuosity of porous media from images by deep learning. Sci Rep 2020;10(1):21488.
[42] Al Khalifah H, Glover PWJ, Lorinczi P. Permeability prediction and diagenesis in tight carbonates using machine learning techniques. Mar Pet Geol 2020;112:104096.
[43] Male F, Jensen JL, Lake LW. Comparison of permeability predictions on cemented sandstones with physics-based and machine learning approaches. J Nat Gas Sci Eng 2020;77:103244.
[44] Wu J, Yin X, Xiao H. Seeing permeability from images: fast prediction with convolutional neural networks. Science Bulletin 2018;63(18):1215–22.
[45] Tembely M, AlSumaiti AM, Alameri WS. Machine and deep learning for estimating the permeability of complex carbonate rock from X-ray micro-computed tomography. Energy Rep 2021;7:1460–72.
[46] Anovitz LM, Cole DR. Characterization and Analysis of Porosity and Pore Structures. Rev Mineral Geochem 2015;80(1):61–164.
[47] Wang F, Yang K, You J, Lei X. Analysis of pore size distribution and fractal dimension in tight sandstone with mercury intrusion porosimetry. Results Phys 2019;13:102283.
[48] Shen R, Zhang X, Ke Y, Xiong W, Guo H, Liu G, et al. An integrated pore size distribution measurement method of small angle neutron scattering and mercury intrusion capillary pressure. Sci Rep 2021;11(1).
[49] Yao Y, Liu D. Comparison of low-field NMR and mercury intrusion porosimetry in characterizing pore size distributions of coals. Fuel 2012;95:152–8.
[50] Schmitt M, Fernandes CP, da Cunha Neto JAB, Wolf FG, dos Santos VSS. Characterization of pore systems in seal rocks using Nitrogen Gas Adsorption combined with Mercury Injection Capillary Pressure techniques. Mar Pet Geol 2013;39(1):138–49.
[51] Li X, Kang Y, Haghighi M. Investigation of pore size distributions of coals with different structures by nuclear magnetic resonance (NMR) and mercury intrusion porosimetry (MIP). Measurement 2018;116:122–8.

[52] Klobes P, Riesemeier H, Meyer K, Goebbels J, Hellmuth KH. Rock porosity determination by combination of X-ray computerized tomography with mercury porosimetry. Fresenius' journal of analytical chemistry 1997;357(5):543–7.
[53] Liu K, Ostadhassan M, Sun L, Zou J, Yuan Y, Gentzis T, et al. A comprehensive pore structure study of the Bakken Shale with SANS, N2 adsorption and mercury intrusion. Fuel 2019;245:274–85.
[54] Basan PB, Lowden BD, Whattler PR, Attard JJ. Pore-size data in petrophysics: a perspective on the measurement of pore geometry. Geological Society, London, Special Publications 1997;122(1):47–67.
[55] Yang X, Mehmani Y, Perkins WA, Pasquali A, Schönherr M, Kim K, et al. Intercomparison of 3D pore-scale flow and solute transport simulation methods. Adv Water Resour 2016;95:176–89.
[56] Yang X, Scheibe TD, Richmond MC, Perkins WA, Vogt SJ, Codd SL, et al. Direct numerical simulation of pore-scale flow in a bead pack: Comparison with magnetic resonance imaging observations. Adv Water Resour 2013;54:228–41.
[57] Oostrom M, Mehmani Y, Romero-Gomez P, Tang Y, Liu H, Yoon H, et al. Pore-scale and continuum simulations of solute transport micromodel benchmark experiments. Comput Geosci 2016;20(4):857–79.
[58] Supervised VV. Learning with Python 2020.
[59] de Myttenaere A, Golden B, Le Grand B, Rossi F. Mean Absolute Percentage Error for regression models. Neurocomputing 2016;192:38–48.
[60] Mohamad IB, Usman D. Standardization and Its Effects on K-Means Clustering Algorithm. Research Journal of Applied Sciences, Engineering and Technology 2013;6(17):3299–303.
[61] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. Journal of machine learning research 2011.
[62] Farquharson JI, Wadsworth FB. Upscaling permeability in anisotropic volcanic systems. J Volcanol Geoth Res 2018;364:35–47.
[63] Warren JE, Price HS. Flow in Heterogeneous Porous Media. Soc Petrol Eng J 1961;1(03):153–69.
[64] Cardwell WT, Parsons RL. Average Permeabilities of Heterogeneous Oil Sands. Transactions of the AIME 1945;160(01):34–42.
[65] Comisky JT, Newsham K, Rushing JA, Blasingame TA. A comparative study of capillary-pressure-based empirical models for estimating absolute permeability in tight gas sands. SPE Annual Technical Conference and Exhibition 2007;3:1821–38. https://doi.org/10.2118/110050-MS.