Characterisation and Modelling of Heterogeneous Sandstone and Carbonate Rocks

Nele Wenck[1], Samuel J. Jackson[2,1], Ann Muggeridge[1], and Samuel Krevor[1]

[1]Department of Earth Science and Engineering, Imperial College London, United Kingdom
[2]CSIRO Energy, Private Bag 10, Clayton South, Victoria 3169, Australia

Key Points:

- Successful capillary heterogeneity characterisation in carbonates depends on the extent to which key features are resolved in X-ray imagery
- Isotropic capillary heterogeneity in carbonates results in a non-monotonic rate-dependant relative permeability
- The work underscores both the importance and the challenges of characterising small-scale heterogeneity for subsurface fluid flow

Corresponding author: Nele Wenck, `nele.wenck15@imperial.ac.uk`

Abstract

The characterisation of multiphase flow properties is essential for predicting large-scale fluid behaviour in the subsurface. Insufficient representation of small-scale heterogeneities has been identified as a major gap in conventional reservoir simulation workflows. Capillary heterogeneity has an important impact on small-scale flow and is one of the leading causes of anisotropy and flow rate dependency in relative permeability. We evaluate the workflow developed by Jackson et al. (2018) for use on rocks with complex heterogeneities. The workflow characterises capillary heterogeneity at the millimetre scale. The method is a numerical history match of a coreflood experiment with the 3D saturation distribution as a matching target and the capillary pressure characteristics as a fitting parameter. Coreflood experimental datasets of five rock cores with distinct heterogeneities were analysed: two sandstones and three carbonates. The sandstones exhibit laminar heterogeneities. The carbonates have isotropic heterogeneities at a range of length scales. We found that the success of the workflow is primarily governed by the extent to which heterogeneous structures are resolved in the X-ray imagery. The performance of the characterisation workflow systematically improved with increasing characteristic length scales of heterogeneities. Using the validated models, we investigated the flow rate dependency of the upscaled relative permeability. The findings showed that the isotropic heterogeneity in the carbonate samples resulted in non-monotonic behaviour; initially the relative permeability increased, and then subsequently decreased with increasing flow rate. The work underscores the importance of capturing small-scale heterogeneities in characterising subsurface fluid flows, as well as the challenges in doing so.

1 Introduction

Subsurface multiphase fluid flow is central to a number of scientific and engineering processes of major societal importance including energy resource use, environmental contaminant remediation, and climate change mitigation. Despite this, there are long-standing difficulties with the characterisation and predictive modelling of subsurface flow. For example, at many CO_2 storage sites worldwide, carbon dioxide injected underground has migrated away from injection points at much faster rates than had been predicted with reservoir simulations (Arts et al., 2004; Chadwick et al., 2009; Hosseini et al., 2013; Lu et al., 2013; Birkholzer et al., 2015; Global CCS Institute, 2019; Onoja & Shariatipour, 2019), see Aminu et al. (2017). These observations suggest that flow simulations are missing some of the leading order mechanisms governing fluid flow.

One gap in conventional reservoir simulation workflows is the field scale representation of the impact of small-scale heterogeneity in multiphase flow properties, the relative permeability and capillary pressure characteristics (Jackson & Krevor, 2020). Relative permeability is one of the key parameters controlling fluid behaviour (S. Krevor et al., 2019). It is strongly influenced by rock heterogeneity that arises from sedimentary structures like crossbedding in sandstones or shell fragments in carbonates (Corey & Rathjens, 1956; Huppler, 1970; Kortekaas, 1985; Hove et al., 1990; Hamon & Roy, 2000; Dawe et al., 1992; Chang & Yortsos, 1992). Field scale simulations typically use relative permeability curves measured in the laboratory using cm-scale rock cores, however capillary pressure heterogeneity has a particularly important impact on flow at these length scales. Capillary heterogeneity is one of leading causes of anisotropy and flow rate dependency in observed relative permeability and has significant impacts on upscaled properties (Woods & Farcas, 2009; Yamamoto, 2009; Green & Ennis-King, 2010; Perrin & Benson, 2010; S. C. Krevor et al., 2011; S. Krevor et al., 2015; Benham et al., 2020; Jackson & Krevor, 2020).

The ratio of viscous to capillary forces over length scales of centimeters to meters controls the importance of capillary heterogeneity. A continuum scale capillary number is often used to describe this such as the number defined by Virnovsky et al. (2004),

$$N_c = \frac{H}{L}\frac{\Delta P}{\Delta P_c}, \tag{1}$$

where H [m] is a length scale associated with the heterogeneity, e.g. a layer thickness, L [m] is a length scale in the direction of flow, ΔP [Pa] is the pressure differential across L, and ΔP_c [Pa] is a contrast in capillary pressure imposed by the heterogeneities. At high capillary numbers the role of heterogeneities diminishes and a single relative permeability characteristic controls the flow in the rock. This is known as the viscous-limit. As capillary forces become significant, heterogeneities control the fluid behaviour and flow rate dependent and anisotropic relative permeability functions are needed to represent the impact of these heterogeneities on large-scale flow. In the subsurface, flow is more prevalently characterised by the capillary-controlled flow regime and capillary heterogeneity is one of the dominant fluid distribution mechanisms over centimeter to meter length scales for oil-brine and gas-brine systems (Chang & Yortsos, 1992; Ringrose et al., 1993; Chaouche et al., 1994; Huang et al., 1995). Hence, upscaled equivalent functions are required to represent the larger scale manifestations of small-scale heterogeneities (Jackson et al., 2018; Jackson & Krevor, 2020).

An approach to characterise capillary heterogeneity within cores has been developed, which combines experimental and numerical methods (Krause et al., 2011, 2013). The workflow uses a numerical history match of a coreflood experiment with the 3D saturation distribution as a matching target and the capillary pressure characteristics as a fitting parameter. The authors applied their method to two sandstone samples and successfully predicted the sub-core scale saturation distribution, with a correlation $R^2 > 0.93$. Subsequently, this approach was applied to a wider range of samples, mainly sand-

stones, and developed to improve the observational basis and strengthen the iterative matching procedure (Berg et al., 2013; Krause et al., 2013; Pini & Benson, 2013a, 2013b; Kong et al., 2015; Krause & Benson, 2015; Jackson et al., 2018; Reynolds et al., 2018; Hosseinzadeh Hejazi et al., 2019).

Uncertainty surrounding the impact of heterogeneity and the resulting relative permeability anisotropy and rate dependency in more complex systems, for instance reservoir sandstones or carbonates, remains. Reservoir sandstones can be significantly affected by diagenic and compactional processes, which lead to substantial porosity and permeability variations (Worden et al., 2018; Heidsiek et al., 2020). These often cluster into complex sedimentary structures such as cross-bedding or flute casts. Carbonate rocks are also characterised by heterogeneity, though on a larger range of scales. On a core to field scale, they exhibit facies and diagenetic distribution patterns, often at sub-seismic resolution (Petrovic et al., 2018). At the pore-scale, variations in pore network topology and wettability add further complexity to spatial porosity and permeability distributions, the latter manifesting as isotropic heterogeneity at larger scales (Rebelle et al., 2009). These complex rock types comprise a major reservoir resource and techniques for rock characterisation capturing the impact of these heterogeneities is important for field studies (Al-Kharusi & Blunt, 2008; Sayers, 2008).

In this work, we evaluate the ability of the history matching workflow described by Jackson et al. (2018) to characterise capillary heterogeneity in rocks with more complex heterogeneities. We apply the method to five reservoir samples with distinct types and length scales of heterogeneity: two sandstones and three carbonates. For the sandstones we use experimental dataset of Reynolds et al. (2018) and for the carbonate rocks, the dataset of Manoorkaar et al. (2021). The sandstones exhibit distinctly-orientated planar bedding, whereas the carbonates are characterised by isotropic cementation of varying length scale. The investigation also allowed us to compare the rate dependency of relative permeability in different rocks and draw conclusions on the varying flow behaviour.

2 Methods

In this section we first describe the rock samples analysed in this study. We then discuss the details of the numerical simulations including the simulator used and the workflow followed to build the numerical models. Lastly, the history match to characterise the capillary heterogeneity is described.

2.1 Rock Samples

Previously acquired experimental datasets using five rock cores were studied, with experimental methods and data reported in Reynolds and Krevor (2015); Reynolds et al. (2018); Manoorkaar et al. (2021). These datasets comprise observations from corefloods performed on two sandstones and three carbonate rock samples. Steady-state experiments were performed with the co-injection of nitrogen and DI water or CO_2 and brine at high (HR) and low (LR) flow rates to obtain flow parameters in the viscous-limit (VL) and capillary-limit (CL) flow regimes, with 3D X-ray images taken throughout using a medical X-ray CT scanner.

The five rock samples cover a range of depositional facies and exhibit distinct heterogeneity types and lengthscales, Figure 1 and 2. The Bentheimer sandstone is a shallow-marine deposit widely used for experimental studies due to its homogeneity (Peksa et al., 2015). Our sample exhibits a simple porosity heterogeneity orientated as a single layer parallel to the flow direction and serves as a nearly homogeneous benchmark. The Bunter sandstone is a geological unit from a previously proposed CCS site in the Southern North Sea. It was deposited in a predominantly alluvial environment and exhibits characteristics of early diagenetic processes such as grain dissolution and cementation (Brook et al., 2003). As such, our sample features noticeable heterogeneity in porosity and permeability, which can be grouped into distinct layers perpendicular to the axis of flow.

The three carbonate samples, Indiana limestone, Estaillades limestone, and Edwards Brown dolomite, exhibit varying degrees of cementation resulting in distinct lengthscales of isotropic heterogeneity. The Indiana limestone, quarried from the Salem Formation in Indiana (USA) is formed of mainly calcite cemented grain stone (El-Maghraby, 2012). Out of the five rock samples, it has the smallest scale porosity heterogeneity ($\approx$ 1mm) and appears relatively homogeneous in the porosity profile. The Estaillades, a calcite-rich limestone, originates from a quarry in Southeast France and is characterised by intergranular macropores and intragranular micropores (Lai et al., 2015; Al-Menhali et al., 2016). The sample has an order of magnitude larger scale of heterogeneity ($\approx$ 1 cm) as shown in Figure 1. The Edwards Brown dolomite, sampled from the Upper Cretaceous formation in Texas (USA) (Manoorkaar et al., 2021), has a multi centimetre length low-porosity region towards the outlet of the core. Table 1 summarises the experimental and modelling parameters for the samples.

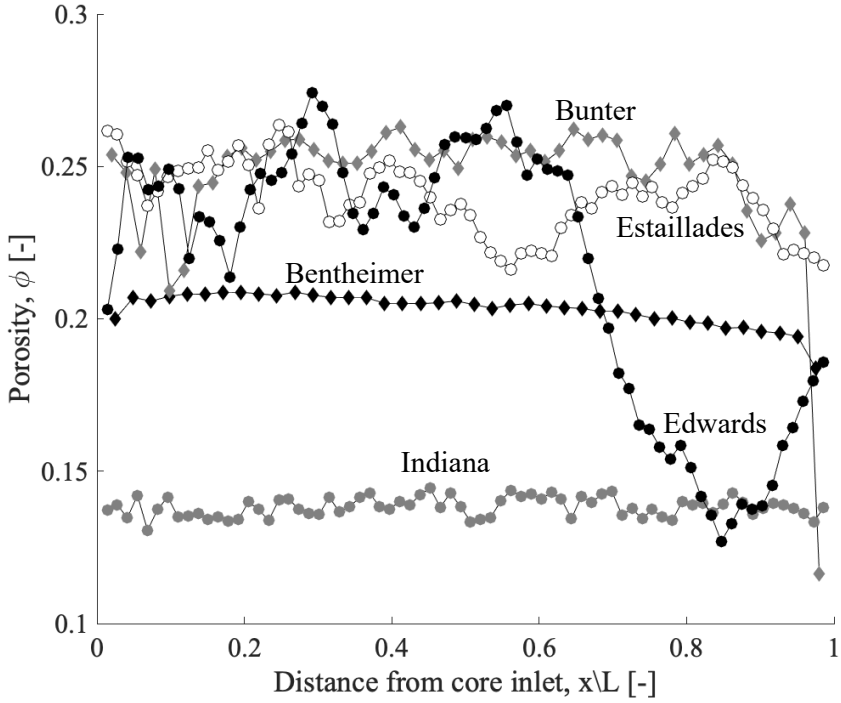

Figure 1. Porosity profiles for the five cores displaying the range of heterogeneity types and length-scales.

Table 1. Summary of the experimental and core characterisation parameters for the five samples.

Parameters	**Bentheimer**	**Bunter**	**Indiana**	**Estaillades**	**Edwards Brown**
Nonwetting phase, nw	N_2	CO_2	N_2	N_2	N_2
Wetting phase, w	DI Water	Brine	DI Water	DI Water	DI Water
Q_{tot} high/low (ml min^{-1})	40/7	20/0.2	0.5/5	0.5/20	0.5/5
Number of fractional flows high/low	10/6	8/6	13/13	13/10	10/16
Pressure, P (MPa)	15.5	13.1	10	10	10
Temperature, T (°C)	50	53	20	20	20
nw density, ρ_{nw} (kg m^{-3})	115	604	115	115	115
w density, ρ_w (kg m^{-3})	997	1022	998	998	998
nw viscosity, μ_{nw} (μPa s)	22.1	45.9	19.7	19.7	19.7
w viscosity, μ_w (μPa s)	550	582	1070	1070	1070
IFT, γ (mNm^{-1})	62	34.7	62	62	62
Experimental core length, L (m)	0.198	0.151	0.148	0.148	0.148
Experimental core radius, r (m)	0.019	0.019	0.019	0.019	0.019
Raw voxel Δx, Δy (m)	0.00023	0.00023	0.00017	0.00017	0.00017
Raw voxel Δz (m)	0.005	0.003	0.001	0.001	0.001
Digital core dimensions [x,y,z]	11x11x41	11x11x52	9x9x74	9x9x74	9x9x74
Upscaled voxel Δx, Δy (m)	0.0032	0.00277	0.00246	0.00246	0.00246
Upscaled voxel Δz (m)	0.005	0.003	0.002	0.002	0.002
Digital core length, L (m)	0.195	0.150	0.144	0.144	0.144
Digital core radius, r (m)	0.018	0.015	0.011	0.011	0.011
Entry pressure, P_e (kPa)	3.51	1.62	2.96	9.30	9.18
Pore distribution, λ (-)	2.3	1.43	0.86	1.45	0.48
Porosity, ϕ (-)	0.21	0.25	0.14	0.24	0.23
Permeability, K_{abs} (D)	1.86	2.20	0.024	0.14	0.046
Chierici k_{rw}, A/L (-)	3/0.75	3/0.9	4.37/1.01	5.26/0.8	15.7/1.06
Chierici k_{rg}, B/M (-)	5/0.65	3.75/0.4	3.30/0.83	1.21/1.22	2.66/0.54
$k_{rg}(S_{wirr}$/ $k_{rw}(S_{gc}$ (-)	1/1	1/1	1/1	1/1	1/1
S_{wirr}, S_{gc} (-)	0.08/0.0	0.082/0.0	0.0/0.0	0.07/0.0	0.0/0.0

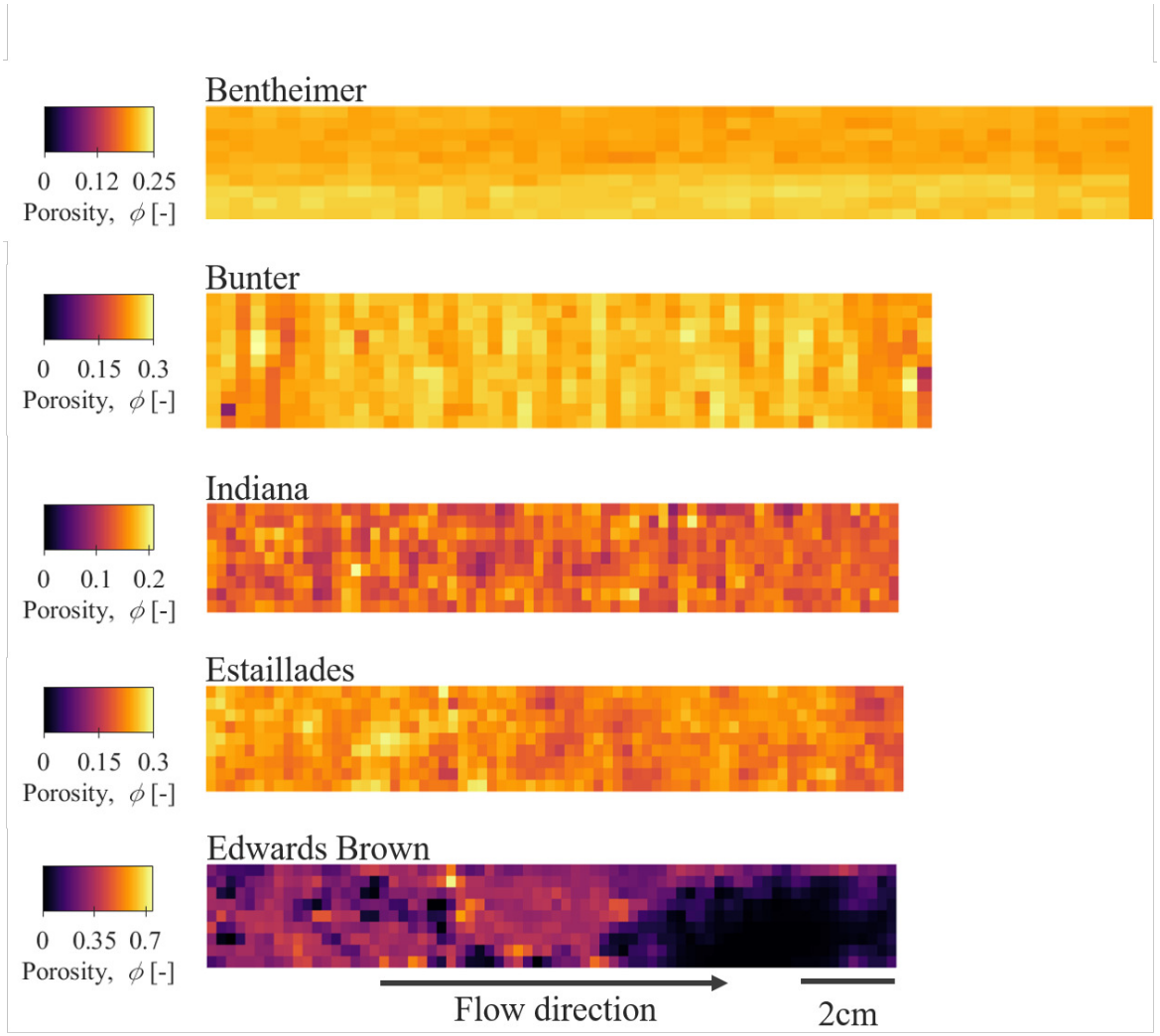

Figure 2. Porosity maps through the centre of each rock sample showing the 2D structure of the porosity heterogeneities.

2.2 Numerical Modelling

The flow simulations were performed using a fully-implicit, isothermal black oil fluid simulator (CMG ™IMEX). The simulator uses the finite difference method to solve the governing equations.

The grid dimensions used in the flow simulations of each core are given in Table 2. The simulations of each core used the rock and fluid properties given in Table 1 together with 3D porosity data obtained from medical CT images, processed following the standard method detailed in Withjack (1988).

The viscous-limit relative permeability data were obtained from data measured during high flow rate experiments. The parameters obtained by Jackson et al. (2018) were used for the sandstone cores whilst the data for the carbonate cores were obtained by history matching using the 1D Simulator SENDRA. The history match used the slice-average saturations, pressure drops, and fluid injection rates. For the Indiana limestone and the Edwards Brown dolomite, SENDRA failed to converge to a solution with an acceptable residual error. Thus, Matlab's "fmincon" function, a gradient-based constrained optimisation tool, was applied to the experimentally measured viscous-limit relative permeabilities. All viscous-limit relative permeabilities were assumed to be uniform throughout the rock domain and were modelled using the Chierici functional form (Chierici, 1984):

$$k_{rg} = k_{rg}(S_{wirr})e^{-BR_w^m}, \qquad k_{rw} = k_{rw}(S_{gc})e^{-AR_w^{-L}}, \qquad R_w = \frac{S_w - S_{wirr}}{1 - S_{gc} - S_w}, \tag{2}$$

where k_{rg} and k_{rw} are the gas and water relative permeabilities, respectively. S_w, S_{wirr} and S_{gc} refer to the water saturation, irreducible water saturation and critical gas saturation, respectively. A, B, M and L are the Chierici parameters that control the shape of the curves.

An average, or intrinsic, capillary pressure characteristic was obtained from observations made using mercury injection porosimetry (MIP). The data was fit with the Brooks-Corey model (Brooks & Corey, 1964):

$$P_c(S_w) = P_e(\frac{1 - S_{wirr}}{S_w - S_{wirr}})^{\frac{1}{\lambda}}, \tag{3}$$

where P_c [Pa] is the capillary pressure as a function of water saturation (S_w [-]), P_e [Pa] is the entry pressure, S_{wirr} [-] is the irreducible water saturation and λ [-] is the pore size distribution factor. The values of P_e, S_{wirr} and λ from Jackson et al. (2018) were used in the sandstone workflow. For the carbonates, the MIP data was cropped prior to fitting due to the large pressures (>1000kPa) reached during mercury injection. Thereafter, the capillary pressure at $S_w = 1.0$ was taken as P_e, and S_{wirr} and λ were obtained by minimising the misfit between the MIP data and Equation 3.

Fictitious inlet and outlet slices were used to mimic the experimental conditions. The following parameters were assigned to these end slices:

1. Linearly varying relative permeabilities: $k_{rw} = S_w$, $k_{rg} = -S_w + 1$
2. Zero capillary pressure: $P_c = 0$
3. A large permeability: $K_{endslice} \approx 7 \times K_{abs}$
4. A constant porosity: $\phi_{endslice} = \phi_{core,avg}$

Injection 'wells' were defined in the centre of the inlet slice — two injected gas and two injected water. Two fluid sink locations – modelled as production wells – were placed

in the centre of the outlet slice. The supplementary material of Jackson et al. (2018) provides a more detailed discussion on the choice of these boundary conditions.

The grid size for the digital models was guided by the representative elementary volume (REV) of the rock structure, the experimental saturation precision and the run time of the simulations. With respect to the concept of REV, the focus of this study was to evaluate the impacts of heterogeneity. Thus we are not identifying a length scale at which an average property can be taken to be representative of the entire core. The relevant issue for this work is to verify that the voxel scale chosen is large enough such that a continuum property has meaning. This can be estimated by comparing the pore structure of each rock type to the voxel size.

For the Indiana limestone, a pore on average has a volume of $22\mu m^3$ (using the median pore-throat radius, r_{50}, from MIP data presented in Lai et al. (2015)). From the voxel dimensions (Table 1) the coarsened voxel size of the CT images is $1.21\times10^{10}\mu m^3$. An estimated $1.69 \times 10^9\mu m^3$ of pores is present within each voxel. This translates to $\approx 7.5\times10^7$ pores. Following the same workflow, the number of pores within each voxel is estimated as $\approx 2.1\times10^8$ and $\approx 3.1\times10^6$ for the Estaillades limestone and Edwards Brown dolomite, respectively.

We note the importance of additionally considering the REV of the capillary pressure characteristic as this might be different from the REV of the pore structure. However, quantifying this poses a challenge. Previous work using micrometre resolution X-ray imaging and pore network models has identified that correlation length scales of capillary pressure characteristics and porosity are of a similar order of magnitude (Jackson et al., 2020; Zahasky et al., 2020). We assume, based on the above estimates, that the voxel scale allows for the capillary pressure characteristic to also be considered a valid continuum property.

The uncertainty in the experimental saturation data was also considered in the numerical modelling approach. We estimated the uncertainty in the saturation values obtained from the experiments following from the analysis presented in Pini et al. (2012). X-ray CT images at centimetre scale resolution are significantly affected by random noise. Averaging data over repeat scans taken during the experiment and processing the images through coarsening can reduce the image noise. Figure S1 in the supplementary material shows the reduction in voxel scale uncertainties from image coarsening and averaging over repeated scans. Table 2 summarises the final saturation uncertainties associated with the coarsening schemes used in this work. The voxel saturation uncertainties for all samples fall below 0.05. We used a grid size of 9×9×74 in the workflow described in Section 2.3.1. Additional coarsening (9×9×37) was applied to speed up the simulations in the extended approach outlined in Section 2.3.2.

Table 2. Uncertainty in the saturations for the five rock samples after specific coarsening schemes were applied to the raw CT images. For the carbonates, 9×9×74 and 9×9×37 are the digital core dimensions used in the capillary heterogeneity characterisation workflow and history match, respectively.

Rock Sample	**Grid size**	$\sigma_{\mathbf{S=0,vox}}$	$\sigma_{\mathbf{S=1,vox}}$
Bentheimer	11×11×41	0.027	0.038
Bunter	11×11×52	0.036	0.051
Indiana	9×9×74	0.011	0.016
	9×9×37	0.008	0.011
Estaillades	9×9×74	0.012	0.001
	9×9×37	0.017	0.012
Edwards Brown	9×9×74	0.020	0.029
	9×9×37	0.014	0.020

2.3 History Match of the Numerical Models

The approach of Jackson et al. (2018) was used to history match the numerical models for all of the experimental datasets. With some of the carbonate rocks, particularly the Indiana and Estaillades limestone, the approach was either unsuccessful or only partially successful. In these cases an extended approach was used, principally characterised by use of the simulation relative permeability property as a tuning parameter, to achieve a satisfactory history match. While this allowed for a more satisfactory match of the experimental data at a specific flow rate condition (either high or low flow rate), it did not result in a model that could predict the impact of varying flow rate away from the conditions at which the model was calibrated. We summarise the approach of Jackson et al. (2018) and the extended approach here.

2.3.1 Approach of Jackson et al. (2018)

The 3D saturation maps of the rock cores obtained during the corefloods were used to infer the heterogeneity in capillary pressure characteristics. It was assumed that the capillary pressure was heterogeneous on a grid block scale. An initial guess was made based on an inversion of the saturation maps, followed by an iterative process whereby capillary pressure characteristics were updated based on the comparison between simulated 3D saturation maps and the observations. The workflow is briefly described below and summarised by the schematic in Figure 3. For further detail, please see Jackson et al. (2018) on which this work was directly based. This in turn built directly on the work of a number of studies: Krause et al. (2011, 2013); Pini and Benson (2013a); Reynolds and Krevor (2015); Reynolds et al. (2018).

For the initial guess, capillary pressure in each slice was assumed to be constant. The average saturation in that slice was assumed to map to the Brooks-Corey fit of the capillary pressure characteristic curve measured during routine core analysis (Equation 3). Voxel scale variation in the saturation within the slice was assumed to be caused by the capillary heterogeneity within the slice. From this a scaling factor κ was assigned to each voxel, adjusting the local capillary pressure characteristic curve, to minimise the mismatch between the slice-average and voxel-specific values:

$$P_{c,ijk}(S_{ijk}) = \kappa_{ijk} \cdot P_{c,avg}(S_{ijk}), \tag{4}$$

$$\Theta = \sum_{i}^{N_v} \sum_{j}^{N_f} \sqrt{(\kappa_{ijk} P_{c,avg}(S_{ijk}^{exp}) - P_{c,ijk}(S_{ijk}^{exp}))^2} \sqrt{S(\kappa_{ijk} P_{c,avg}) - S_{ijk}^{exp}}, \tag{5}$$

where $P_{c,ijk}$ is the individual voxel capillary pressure, $P_{c,avg}$ is the average capillary pressure curve, S_{ijk}^{exp} is the experimental voxel saturation, κ_{ijk} is the individual voxel scaling parameter, N_v is the total number of voxels, N_f is the total number of fractional flows and $S(\kappa_{ijk} P_{c,avg})$ represents the saturation of the average capillary pressure curve after it has been scaled (using the slice-average capillary pressure). Through this, a 3D map of the initial scaling factor was built. The scaling was then used to populate the capillary pressure characteristics of numerical simulations of corefloods using the CMG ™IMEX fluid simulator. The capillary heterogeneity was introduced alongside the porosity profile, the core-average characteristic capillary pressure behaviour and the viscous-limit relative permeability.

The iterative calibration of the capillary pressure characteristics followed from this initial guess. After the first simulation, 3D saturation and capillary pressure maps were

extracted and directly compared to the experimentally measured values. A deviation from the experimental observations (both P_c and S_w) was assumed to stem from an incorrect scaling parameter assigned to the voxel. The scaling factor κ was then updated, minimising the mismatch between the experiment and simulation. The objective function becomes the following, where the slice-average P_c values have been replaced by the simulation values:

$$\Theta = \sum_{i}^{N_v} \sum_{j}^{N_f} \sqrt{(\kappa_{ijk} P_{c,avg}(S_{ijk}^{exp}) - P_{c,ijk}(S_{ijk}^{sim}))^2} \sqrt{S(\kappa_{ijk} P_{c,avg}) - S_{ijk}^{exp}}, \tag{6}$$

where S_{ijk}^{exp} has been replaced with the simulation voxel saturation, S_{ijk}^{sim}. Now, $S(\kappa_{ijk} P_{c,avg})$ represents the saturation of the average capillary pressure curve after it has been scaled using the individual simulated voxel capillary pressure rather than the experimental slice-average capillary pressure. Please see Figure 3 for a diagram summarising this workflow.

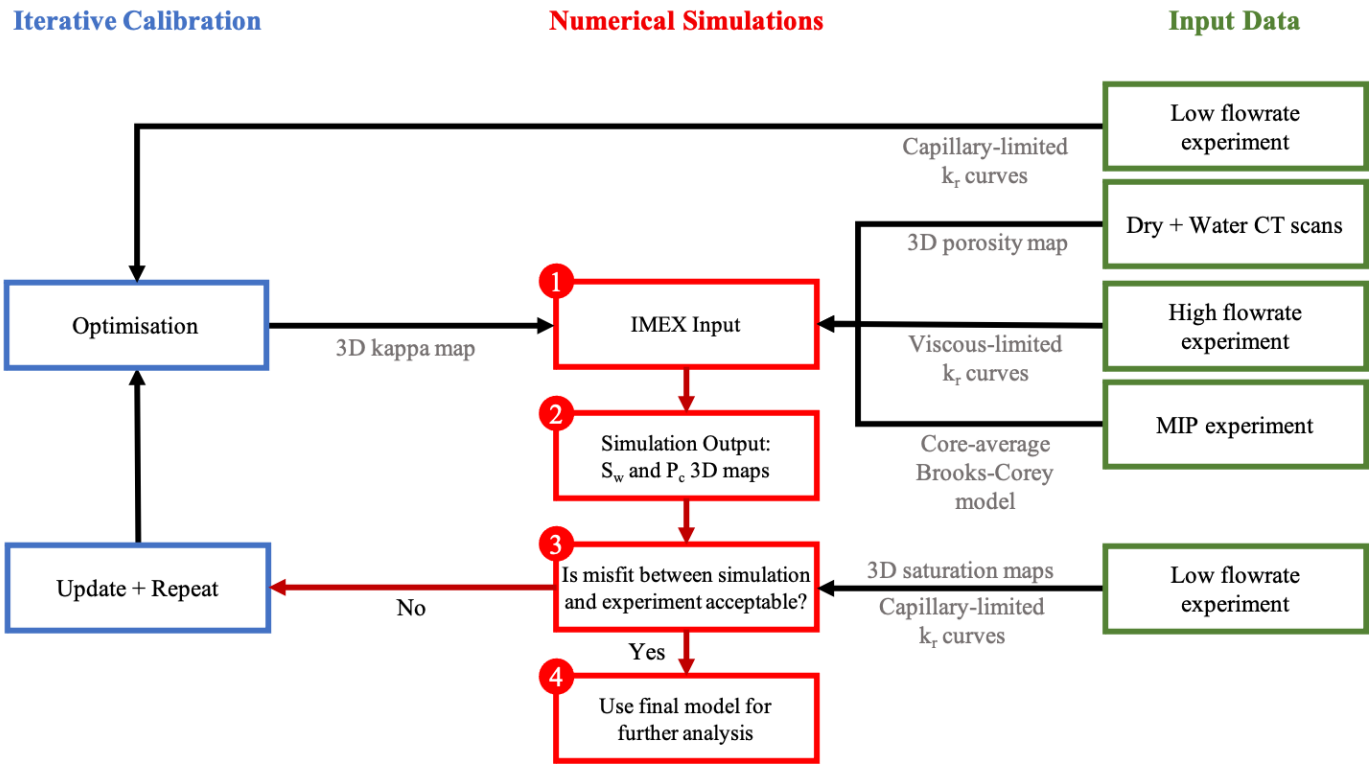

Figure 3. Flowchart summarising the workflow and iterative calibration scheme followed to characterise the capillary heterogeneity in a core.

The iterative calibration scheme is said to have converged when the error between the experiment and simulation voxel saturations stabilised i.e. when the relative change in R^2 between iterations falls below 2%. As shown in Figure 4, the sandstones generally exhibit an earlier convergence compared with the carbonates. The R^2 stabilised after three iterations. The error in the voxel saturations reaches a plateau after four iterations for the Estaillades and Edwards Brown, whereas the Indiana is associated with convergence issues. During the calibration scheme, the κ values assigned to the Indiana core were changed significantly to minimise the error in the voxel saturations. With each

iteration, the range of κ values increased. Beyond five iterations, the κ values reached physically unrealistic values ($\kappa > 700$). Thus, we decided to terminate the optimisation scheme after five iterations as displayed in Figure 4. This issue is discussed in more detail in Section 3.2.1. The R^2 associated with the final calibrated models are summarised in Table 3.

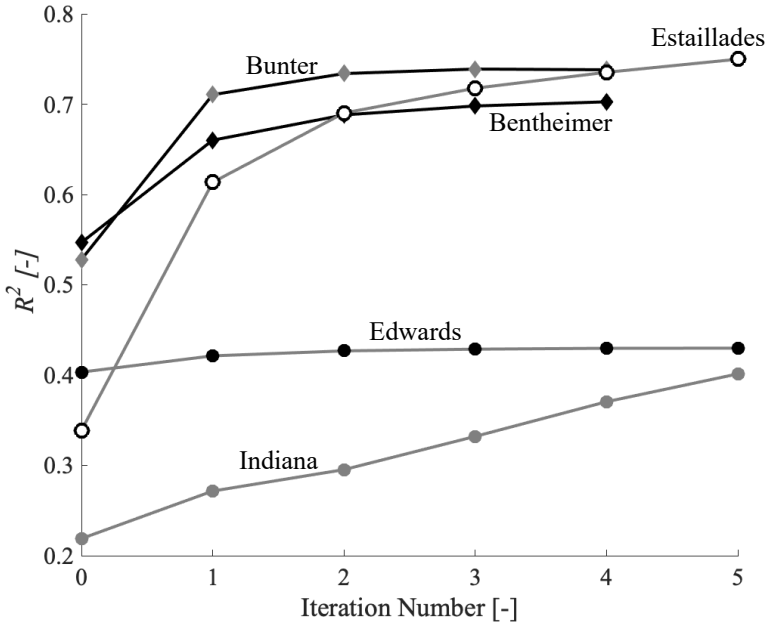

Figure 4. R^2 values for the experiment and simulation voxel saturations for the five rock samples plotted against the number of simulation runs. The sandstones generally exhibit an earlier convergence compared to the carbonates.

Table 3. R^2 values associated with the final calibrated models for each rock type. N_r stands for the number of iterations.

Rock Sample	$\mathbf{N_r}$	$\mathbf{R^2}$
Bentheimer	4	0.701
Bunter	4	0.738
Indiana	5	0.401
Estaillades	5	0.750
Edwards Brown	5	0.430

2.3.2 Extended Approaches Using Relative Permeability as a Matching Parameter

The viscous-limit relative permeability curves used as input to the simulation are a major control on the results and ideally are measured with very little uncertainty in the observation. However, coreflood experiments are challenging and the measurements can be affected by sources of error. With respect to the high rate experiment, it may be difficult for true viscous-limit flow conditions to be achieved in the laboratory for anything other than permeable, relatively homogeneous cores. For highly heterogeneous rocks the highest flow rate achievable may result in a relative permeability measurement still impacted by capillary heterogeneity (Krause & Benson, 2015). This means that it may not serve as a suitable input for the predictive simulation. Furthermore, as discussed in Berg et al. (2021), the extraction of relative permeability from coreflood experiments using Darcy's Law corresponds to significant non-uniqueness and uncertainty. In this work, possibly related to this issue, the observations from the carbonate rocks were difficult to match following the conventional workflow. Thus we explored improvement in the history-matched simulation by additionally using the core-averaged input viscous-limit relative permeability as a fitting parameter, with the observed relative permeability at the low flow rate as an additional matching target. The modified workflow is summarised by the following steps:

1. Initially, a large range of Chierici parameters, A and B, was explored, centred around the Chierici values obtained from the high rate experimental dataset.
2. The viscous-limit input curve was modelled with the new Chierici parameters and the iterative optimisation procedure was followed. Four iterations were run for each parameter combination to obtain the calibrated, digital 3D model.
3. Thereafter, this model was used in simulations of the low rate experiment. The root mean square errors for the relative permeabilities, core-average saturations and voxel saturations were calculated for each dataset.
4. Using the errors from step 3), the parameter space was constrained. The history match was repeated with a smaller range of A and B values.
5. These steps were repeated until an optimal solution was found that minimised the error in the simulated and experimental observations at the low flow rate.

This workflow, which used the observed relative permeability at the low flow rate as an additional matching target, resulted in an unsatisfactory prediction of the voxel saturations and relative permeabilities at high rate. Therefore, we performed an additional history match for each carbonate sample, which also incorporated the high rate experimental data. The procedure used the errors in both, the low and high rate relative permeabilities and voxel saturations, to guide the history match. The updated workflow followed these steps:

1. Initially, a large range of Chierici parameters, A and B, was explored. Each combination of A and B was subsequently tested.
2. The viscous-limit input curve was modelled with the new Chierici parameters and the iterative optimisation procedure was followed. Four iterations were run for each parameter combination to obtain the calibrated, digital 3D model.
3. Thereafter, this model was used in simulations of the low rate and high rate experiments. The root mean square errors for the relative permeabilities, core-average saturations and voxel saturations were calculated for the low and high rate datasets.
4. Using the errors from step 3), the parameter space was constrained. The history match was repeated with a smaller range of A and B values.
5. These steps were repeated until an optimal solution was found that minimised the mismatch between the simulated and experimental observations at the low and high flow rates.

3 Results and Discussion

To evaluate the matching procedure, the voxel-scale fluid saturations were compared between the experiments and the simulations. In the initial approach, the high flow rate, viscous-limit relative permeability was input into the simulations. The evaluation of the predictive capability was based on the predicted observations at the low flow rate, which reflect the impact of the heterogeneities inverted from the experimental saturation data. In the extended numerical approach, with the carbonate rocks, the input relative permeability was used as a tuning parameter for matching the experimental data. Thus, rather than predicting the low rate relative permeabilities, they were used as an additional matching target.

3.1 Bentheimer and Bunter Sandstone Rocks

Following the workflow from Section 2.3.1, digital models for the two sandstones were generated, Figure 5. The predictive capability is assessed in Figure 6. The relative permeabilities from the simulations align well with the experimental data for both samples, although the Bentheimer generally exhibits a closer match. In the Bentheimer sample, the gas relative permeability is raised relative to the viscous-limit curve, whereas the water relative permeability is lowered. This is as expected: the parallel layering allows for the two phases to arrange themselves optimally, discussed in previous work such as Krause and Benson (2015) and Rabinovich et al. (2016). In comparison to this, both relative permeabilities are lowered relative to the viscous-limit curves in the Bunter sandstone. This indicates that the perpendicular layering restricts the flow of the phases. These results also demonstrate the replicability of the digital models: using identical modelling parameters in the unaltered iterative calibration workflow allowed us to regenerate the findings presented in Jackson et al. (2018).

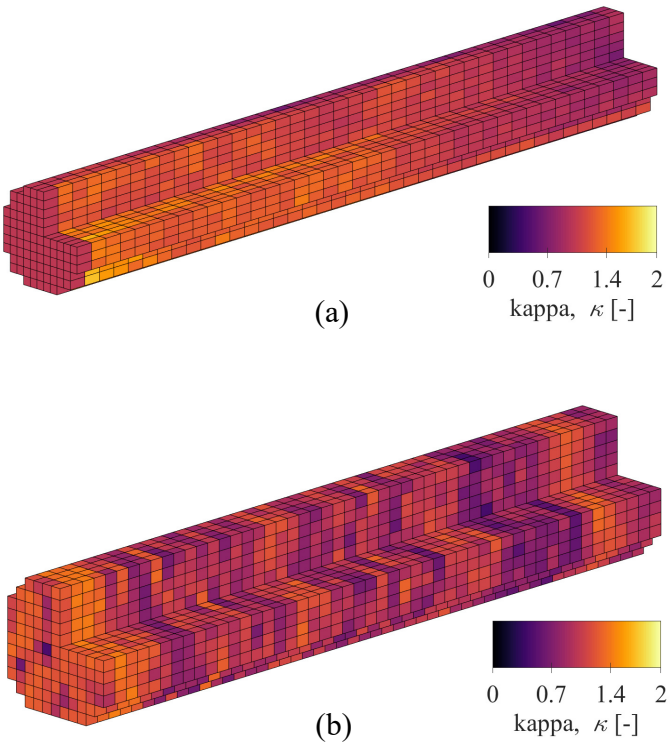

Figure 5. 3D digital model displaying the kappa values obtained after 4 iterations for (a) Bentheimer sandstone and (b) Bunter sandstone. The Bentheimer sandstone exhibits layers parallel to the axis of flow, whereas the Bunter sandstone is characterised by heterogeneities aligned perpendicular to the axis of flow. Digital core dimensions for the Bentheimer and Bunter sandstones respectively: $L = 0.195$m, r= 0.018m and $L = 0.150$m, r= 0.015m.

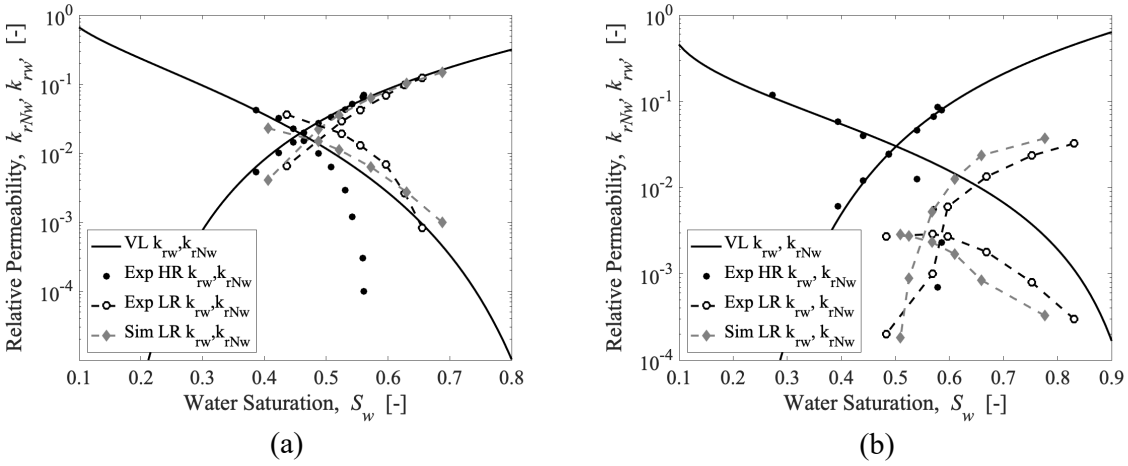

Figure 6. Computed relative permeabilities for (a) the Bentheimer and (b) Bunter sandstone samples obtained from numerical simulations using the iteratively calibrated digital models as input. As shown, the relative permeabilities from the simulations align well with the experimental data for both samples. VL stands for the viscous-limit relative permeabilities used as input to the simulations. Exp LR and HR refer to the low and high rate relative permeabilities respectively, obtained from the coreflood experiment. Sim LR refers to the relative permeabilities obtained from the simulation of the low rate experiment.

3.2 Carbonate Rocks

Numerical models incorporating capillary heterogeneity were created for the carbonates, Figure 7. The matching procedure and predictive capability using the conventional approach, where the experimental viscous-limit relative permeability is constrained, are assessed using Figure 8. The results, when the viscous-limit relative permeability is used as a fitting parameter and the low rate relative permeabilities are an additional matching target, are presented in Figure 9 and Table 4. Each sample is discussed individually (Sections 3.2.1 to 3.2.3). Within each section we first discuss the match using the conventional approach followed by a discussion of the results when the viscous-limit relative permeability is used as a fitting parameter.

Table 4. Comparison of the viscous-limit relative permeability Chierici parameters, A and B, as obtained from the high rate experiment (Sendra was used for the Estaillades, Matlab's "fmincon" tool was used for the Indiana and the Edwards Brown), compared to the results from the history match.

Name	Source of parameters	A	B
Indiana	HR experiment	4.37	3.30
	History match	1.00	0.45
Estaillades	HR experiment	5.26	1.21
	History match	4.75	0.70
Edwards Brown	HR experiment	15.7	2.66
	History match	14.00	2.00

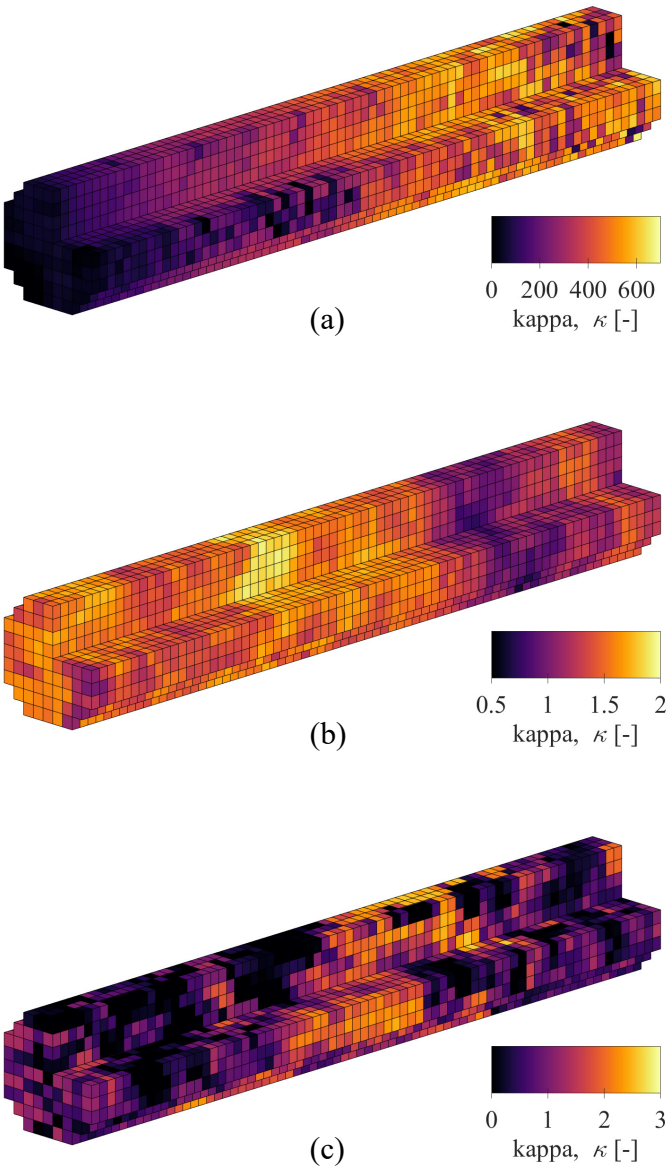

Figure 7. 3D digital model displaying the values of κ obtained after 5 iterations for (a) Indiana limestone, (b) Estaillades limestone and (c) Edwards Brown dolomite. The characteristic length scale of the capillary heterogeneities significantly differs between the samples, where the Indiana limestone is characterised by the finest scaled variations in κ. Digital core dimensions: L = 0.144m, r= 0.011m.

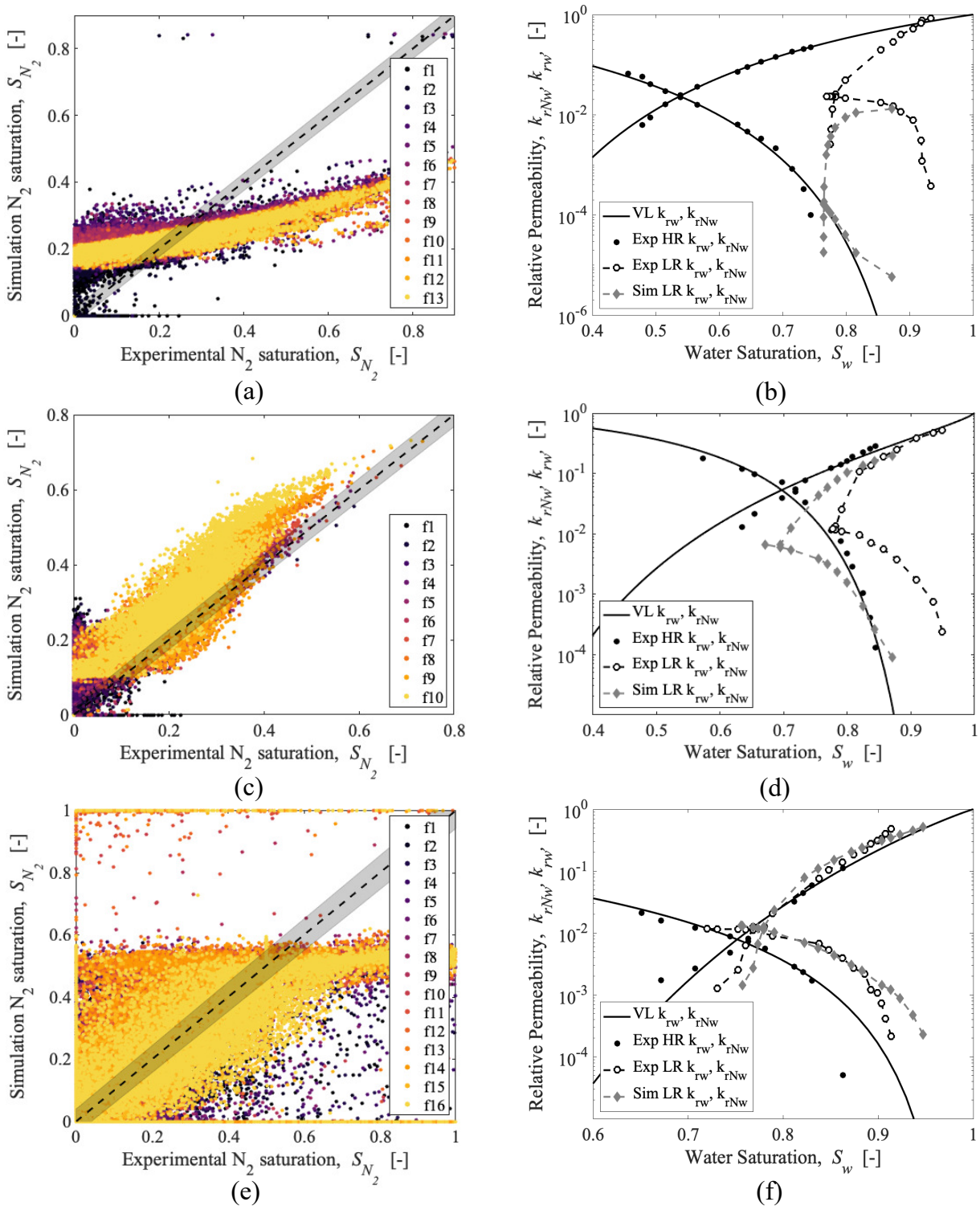

Figure 8. An evaluation of the history match and predictive ability of models of carbonate rocks generated with the high flow rate relative permeability curve as input after 5 iterations. Left: voxel saturation correlation plot comparing the experiment and the simulation for the (a) Indiana, (c) Estaillades and (e) Edwards Brown. The colours correspond to the individual fractional flows. Right: Computed relative permeabilities obtained from numerical simulations using the iteratively calibrated digital models as input and compared to the experimental measurements for the (b) Indiana, (d) Estaillades and (f) Edwards Brown. The Edwards Brown and Indiana are associated with the best and worst predictions of the experimental relative permeability, respectively. VL stands for the viscous-limit relative permeabilities used as input to the simulations. Exp LR and HR refer to the low and high rate relative permeabilities respectively, obtained from the coreflood experiment. Sim LR refers to the relative permeabilities obtained from the simulation of the low rate experiment.

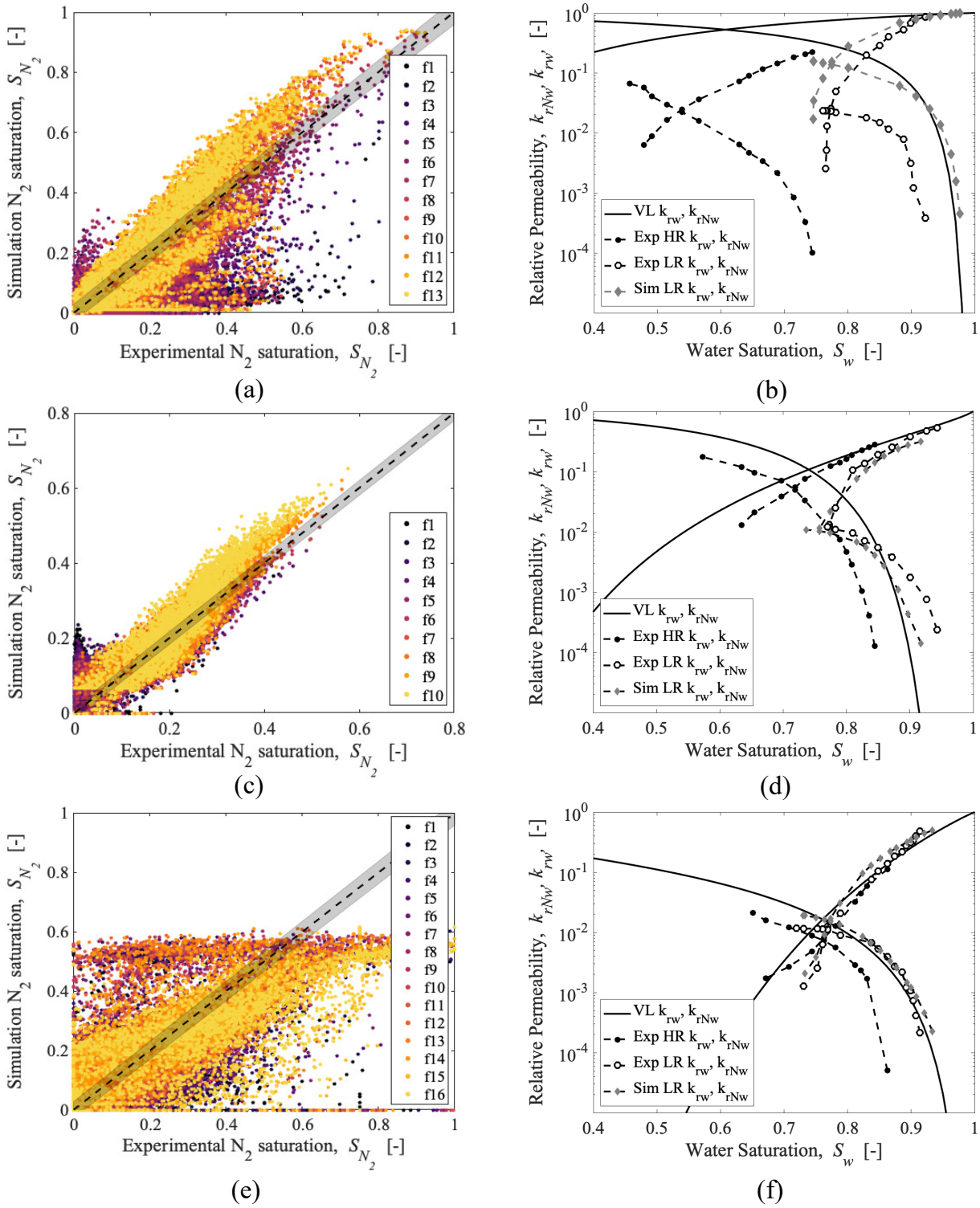

Figure 9. History match results for the carbonates when the viscous-limit relative permeability was additionally used as a fitting parameter and the observed relative permeability at low flow rate was an additional matching target after 4 iterations. Left: voxel saturation correlation plot comparing the experiment and the simulation. The colours correspond to the individual fractional flows. Right: Computed relative permeabilities obtained from numerical simulations using the iteratively calibrated digital models as input and compared to the experimental measurements. Top: Indiana, Middle: Estaillades, Bottom: Edwards Brown. The adapted workflow resulted in a closer match of the voxel saturations and a better prediction of the experimental relative permeability for all samples. VL stands for the viscous-limit relative permeabilities used as input to the simulations. Exp LR and HR refer to the low and high rate relative permeabilities respectively, obtained from the coreflood experiment. Sim LR refers to the relative permeabilities obtained from the simulation of the low rate experiment.

3.2.1 Indiana Limestone

As shown in Figure 7, the Indiana limestone displays finer-scaled variations in κ compared to the Estaillades limestone. Overall, this sample exhibits a large range in entry pressures, where some regions have entry pressures 600 times larger than the core-average. The Indiana limestone has the smallest scale of heterogeneity ($\approx$ 1mm) and it displays the poorest match (Figure 8a). The simulation results in a homogeneous saturation distribution in comparison to the experiment. The majority of voxels have gas saturations $S_{N_2} < 0.4$, which is reflected by the horizontal trend in the voxel scatter plot. In the experiment, the core exhibited a heterogeneous saturation distribution where small regions were associated with anomalously high gas saturations. These are not reproduced in the simulation as can be seen by the significant deviation at high saturations ($S_{N_2} > 0.4$). This is also shown in the 3D saturation maps, Figure S2 in the supplementary material.

This mismatch results in a poor prediction of the relative permeability (Figure 8b). Both, k_{rg} and k_{rw} are underpredicted. The experimental k_{rw} curve has almost 100 times larger relative permeability compared with the simulation. The simulated k_{rg} curve plots closely to the input viscous-limit curve. This suggests that we did not capture key features of the heterogeneity in the imagery, which was responsible for the shift in the relative permeabilities observed in the experiment with varying flow rates. A large core-average pressure drop in the simulation is responsible for the low k_{rg}. This could have been caused by the connectivity of the heterogeneity. If some areas of the core are assigned high P_e, they inhibit the flow and cause a large pressure drop. The κ distribution supports this hypothesis: the optimisation resulted in a large range of κ values, with some reaching nearly 700 (i.e. P_e is 700 times the 2.96kPa core-average), which would significantly hinder fluid invasion into these parts of the core. Comparing different iterations shows that the range of κ values significantly increases with each iteration: from 0-8 to 0-280 between iteration 1 and iteration 4. This large change in κ barely had an impact on the trend produced by the saturation correlation plot: it remained horizontal. This shows that even a significant change in the entry pressure distribution did not affect the outcome. Hence, the iterative optimisation fails to match the experimental observations.

Conducting the history match using the viscous-limit relative permeabilities as fitting parameters resulted in a significant improvement in the match of the low flow rate saturation data, Figure 9a. The capillary-limit relative permeabilities plot notably closer to the experimental data, Figure 9b. Additionally, the voxel saturation correlation follows a linear trend, which suggests the iterative matching was successful using the updated viscous-limit curves. However, the updated models are now unable to predict the saturations or relative permeabilities measured during the high rate experiment. Additionally, the history match produced a model associated with close to no rate dependency: the updated viscous-limit curves (solid black lines) plot very near the simulated low rate relative permeabilities (dashed grey lines). Thus, the shift observed in the experiment with changing flow rate is not reproduced in the history matched simulation. It is thus possible that, unlike the sandstone rock samples, features of the Indiana limestone that are controlling the multiphase flow properties are not resolved in the imagery. Therefore, while we can obtain a satisfactory match in the fluid saturation distribution, this inversion does not result in a physically representative model of the rock core. Motivated by the unsatisfactory prediction of the high rate relative permeabilities, the history match was repeated, this time letting the error of the low and high rate datasets guide the match. However, we were unable to obtain a combination of Chierici parameters that minimised the errors in both datasets. A figure illustrating this can be found in the supplementary material. These findings suggest that there are features in the rock core that have not been sufficiently resolved in the imagery that control the fluid distribution and core-average relative permeability.

3.2.2 Estaillades Limestone

The length scale of heterogeneity for the Estaillades Limestone is between that of the Indiana limestone, with smaller heterogeneities, and the Edwards, which has multi-centimetre scale heterogeneity. The Estaillades limestone is associated with the least variation in entry pressures with a range $0.5 < \kappa < 0.2$, Figure 7b. It exhibits a gradual distribution of κ and these can be clustered into relatively large regions within the core.

The voxel saturation plot displays a linear correlation between the experiment and the simulation (Figure 8c). The iterative calibration partially matches the saturation distribution captured in the CT scans, although they are systematically over estimated. Analysing 3D saturation maps of specific fractional flows underpins this: the saturation distribution pattern is well-reproduced, with a high-saturated region close to the inlet and a low gas saturation two-thirds into the core (see Figure S3 in the supplementary materials). The digital model fails to predict the experimental low rate relative permeabilities (Figure 8d): k_{rg} is significantly underpredicted. The prediction of water relative permeability fits the experimental data better, but the curve is shifted to lower water saturations. This is consistent with the voxel saturation correlations. The experimental data displays a raised k_{rg} relative to the viscous-limit k_{rg}, suggesting that the heterogeneity enhances gas flow. This is not reproduced in the simulation. The offset in k_{rg} is not caused by some areas of the core being assigned very high P_e, the range of κ values is small. The sensitivity to boundary conditions was also tested by setting P_c in the end slice to a finite value. However, the results were unaffected by this. For a detailed discussion, see Section S4 in the supplementary material.

The history match using the input relative permeability as a fitting parameter was applied to the Estaillades dataset. As shown in Figure 9c and d this resulted in a significant improvement in the match to both the saturation and low flow rate relative permeability. However, similarly to the Indiana limestone, the predictive capability of the model was lost. It does not reproduce the observations made at high flow rate. When using both high and low flow rate experimental datasets in the workflow, the history match failed to find a suitable combination of Chierici parameters that minimised the errors in both datasets, similarly to the Indiana limestone. This suggests that the optimisation procedure successfully calibrates to the saturation map and relative permeability at a given flow rate, but fails to result in a physical representation of the rock sample that can replicate the rate dependency of the relative permeability observed in the laboratory.

3.2.3 Edwards Brown Dolomite

The initial matching procedure results in a relatively large spread of voxel saturations, implying that the workflow is not able to match individual voxel saturations well (Figure 8e). However, the spatial distribution is reproduced - a high number of voxels plot closely to the 1:1 trendline. This is also observed in 3D saturation maps, Figure S4 in the supplementary material. The numerical model accurately predicts the capillary-limit relative permeabilities, significantly better than the other two carbonate samples. Overall, this suggests that the model captures the controlling heterogeneity within the Edwards Brown dolomite. The low rate experiment exhibited a raised k_{rg} compared with the viscous-limit curve, which is reproduced in the simulation.

To improve the voxel saturation correlation, we applied the history match workflow to this dataset (Figure 9e and f). Similarly to the Indiana limestone and Estaillades limestone, whilst the match in the voxel saturations significantly improved, the predictive capability of the high rate experiment was lost. Additionally, the shift in the capillary-limit relative permeabilities from the viscous-limit curves observed experimentally, also diminished - the low rate relative permeabilities now plot near the viscous-limit input curves. Once incorporating both datasets into the numerical approach, the workflow failed

to minimise the errors in both datasets, as was observed for the other two carbonate cores. With the Indiana and Estaillades, the use of the viscous-limit relative permeability did not improve the predictive ability of the numerical model. A notable finding with the Edwards is that using the simulation relative permeability as a tuning parameter can lead directly to the loss of predictive capability in the simulation.

3.3 Evaluation of the Workflow Applicability

From the analysis presented in Section 3.2, it emerges that the Indiana limestone resulted in the worst voxel saturation correlation and prediction of the relative permeability. By testing the boundary conditions of the numerical simulations, we ruled out the end-effect as the root cause of the mismatch. Instead, we found that the workflow severely broke down when applied to this sample and the optimisation failed to assign κ values that replicate the experimental $S_w - P_c$ observations. In comparison to this, the Estaillades limestone displayed a good match in the voxel saturations with an R^2 of 0.75. This demonstrates that the workflow successfully calibrated the assigned κ values. However, the digital model failed to predict the experimental relative permeability measured at low flow rate. Instead, the model resulted in notably reduced gas connectivity. The voxel saturation correlation for the Edwards Brown was poor ($R^2 = 0.43$) suggesting that the workflow failed to replicate the voxel-based observations. However, the digital model successfully predicted the low rate relative permeability and the trend in the rate dependency observed experimentally was reproduced. We explored various possibilities responsible for the breakdown of the workflow, which are discussed in the following. First, we focus on the assumptions inherent to the routine rock characterisation, we then evaluate the optimisation and history matching procedures and lastly discuss scaling-related issues.

3.3.1 Complex Pore Structure

The workflow developed by Jackson et al. (2018) models the capillary pressure characteristics using the Brooks-Corey model. Carbonates are associated with complex, often bimodal, pore systems and the flow behaviour significantly changes when transitioning from the invasion of macro to micropores (Cantrell & Hagerty, 1999, 1999; Prodanović et al., 2015). The Brooks-Corey functional form is unable to closely model a bimodal pore distribution, thus could be a source for the mismatch observed in the carbonate samples. Particularly the Estaillades limestone exhibited a significant amount of microporosity, hence became the focus. The Estaillades MIP data was re-fitted using functional forms, which follow the observed bimodal trend more closely. Thereafter the iterative calibration was repeated and the results were compared. However, no visible impact was observed, suggesting the source of mismatch lies elsewhere.

The modelling of the viscous-limit relative permeabilities is also associated with uncertainties. Applying the fully numerical approach to the carbonates (Section 2.3.2) eliminated some factors as the root causes of the mismatch. For instance, the assumption of reaching viscous flow conditions during the high rate corefloods. However, another important consideration is the use of only a single set of viscous-limit relative permeabilities. Microporosity, exhibiting flow behaviour distinct from other regions within a rock, could be characterised by a different set of relative permeabilities. To explore this, more detailed data collection is required. For instance, small plugs corresponding to distinct facies could be cut from core samples. Coreflood experiments could be performed on each of those small samples to infer multiphase flow characteristics associated with a particular facies.

The permeability of the rock core was inferred using Leverett J-function scaling. As discussed in Sarwaruddin et al. (2001), the Leverett J-function is dependant on additional factors including the pore size distribution and the irreducible water saturation. Thus, scaling of the capillary pressure following this method should be limited to rock types displaying a homogeneous pore structure (Sarwaruddin et al., 2001). For the reasons discussed previously, these parameters are likely non-uniform in a carbonate core sample, suggesting that this scaling method is unsuitable. This could have been a factor causing the mismatch between the experiment and simulation observations in the carbonates. To infer the permeability in complex pore structures such as carbonates, pore-based modelling, for instance Lattice Boltzmann or positron emission tomography (PET) methods (Zahasky & Benson, 2018), should be used instead.

The complex pore structure in the carbonates could also enable combined drainage and imbibition to occur during the drainage coreflood experiments. In that case, basing the modelling effort solely on drainage behaviour would not suffice. To explore this aspect further, carbonate coreflood experimental observations at the pore scale would be required.

3.3.2 Optimisation and History Match

Non-uniqueness is a well-known issue in history matching and so we evaluated whether this was a source of uncertainty in the workflow and could explain the poor matches to some of the experimental data. Rather than converging to a global minimum, the optimisation procedure could have converged to a local minimum. By using multiple datasets (low and high flow rate experimental data) and $S_w{-}P_c$ observations from all fractional flows, the simulations were better constrained, reducing the potential impact of non-uniqueness. A further indication that non-uniqueness was unlikely to be the cause of the poor matches was the good matches obtained to the sandstone observations.

Grid convergence issues could also have impacted the simulation outcomes. To evaluate this, a locally-refined grid was implemented in the simulation of one of the carbonate samples. The resultant voxel saturation correlation displayed no improvement, hence grid convergence was eliminated as a major control on the modelling results.

At this stage, it is also worth comparing the method developed by Jackson et al. (2018) with previous heterogeneity characterisation studies. The workflow applied herein uses coreflood datasets obtained at two distinct flow rates, which provide detail on the flow rate dependency by covering a large range of experimental conditions. Furthermore, $S_w{-}P_c$ from all fractional flows (max. 16 for the Edwards Brown) are incorporated into the calibration effort of the capillary pressure heterogeneity. Thus, the optimisation and characterisation is guided by high volume of data to ensure the physical properties of the rock sample are replicated. In comparison to this, previous workflows were primarily based on observations made at one flow rate, and often incorporated $S_w{-}P_c$ voxel measurements from one fractional flow (Krause et al., 2011; Hosseinzadeh Hejazi et al., 2019; Ni et al., 2019). While this results in a close match between the experiment and simulation observations in a distinct flow regime, it likely does not provide sufficient characterisation of flow rate dependency and would thus fail to predict relative permeabilities at arbitrary flow rates.

3.3.3 Scaling

The success of the workflow correlates with the characteristic length scale of the heterogeneity in each sample. The Indiana limestone with mm-large heterogeneity led to the poorest result. This is followed by the Estaillades limestone, which is characterised by cm-large porosity variations. The Edwards Brown dolomite was the only carbonate sample that resulted in a digital model with good predictive capability - its heterogene-

ity was clearly resolved by the imagery. Thus, a factor likely responsible for the breakdown of the calibration is the extent to which key features are resolved by the medical CT scanner. For the Indiana and Estaillades limestone samples, the CT images fail to provide sufficient detail to infer the capillary heterogeneity, which ultimately results in poor voxel saturation correlations and unsuccessful predictions of the relative permeabilities.

Furthermore, the small-scale variations in permeability and porosity, as exhibited by the Indiana and Estaillades limestone samples, could have formed large capillary pressure gradients over short length scales. Consequently, the assumption of capillary pressure equilibrium, which formed the basis of the characterisation workflow, would have broken down. This also could have significantly impacted the modelling of these two samples.

3.4 Rate-Dependant Behaviour of Relative Permeability

Using the iteratively calibrated models for the Bentheimer sandstone, Bunter sandstone and Edwards Brown dolomite, simulations can be run at varying flow rates to investigate the rate dependency of relative permeability in detail. As the workflow failed to sufficiently characterise the heterogeneity in the Indiana and Estaillades limestone samples, they were disregarded. Conducting the rate dependency analysis using numerical simulations allows us to remove laboratory constraints; it is possible to model a wider range of flow rates numerically than can be examined in the laboratory.

To quantify the relative importance of viscous and capillary forces for each of the total flow rates used, the capillary number (Equation 1) was calculated for each experiment. For a detailed discussion on the values used in Equation 1, see Jackson et al. (2018) and Manoorkaar et al. (2021) for the sandstones and carbonates, respectively. Similarly to Jackson et al. (2018), we used the pressure drop at $f_{N_2} = 0.5$ for the calculations. The results are shown in Figure 10.

The trend in the rate dependency varies strongly between the sandstones and the carbonates. The perpendicular layering in the Bunter sandstone generally reduces the permeabilities of both phases, and as the total rate is increased, the relative permeabilities also increase. There is a clear relationship between total rate and resultant permeability, which is indicated by the colour shading of the lines in Figure 10b. In comparison to this, the Bentheimer sandstone, with parallel layering, allows for the phases to align optimally, leading to a raised gas relative permeability relative to the viscous-limit curve. This can be partly seen in Figure 10a.

The carbonate sample displays a different trend. Of key interest is that, distinct from the sandstones, the variation in relative permeability is non-monotonic with varying flow rate. From the lowest flow rate, the relative permeability initially increases in the wetting and non-wetting fluid phases before decreasing again towards the viscous-limit curve. This is related to the nature of the heterogeneity: rather than exhibiting planar bedding, the Edwards Brown is characterised by an isotropic heterogeneity. This non monotonic behaviour was hypothesised in the simulations of Virnovsky et al. (2004). Additionally, the observed rate dependency is weaker than in the sandstones. This is because of the size of the heterogeneity: the Edwards Brown is characterised by a large cemented region. Ultimately, this controls the fluid behaviour even in the high rate experiment. At varying rates, the shift in relative permeabilities is stronger for k_{rw} than k_{rg}. For example, the simulation at $q_{tot} = 0.05$ ml/min results in a significant reduction in the water relative permeability (darkest-shaded line), whereas the gas relative permeability plots closely to the viscous-limit curve. This suggests that the relative strength of capillary forces has a stronger impact on the water flow compared with the gas.

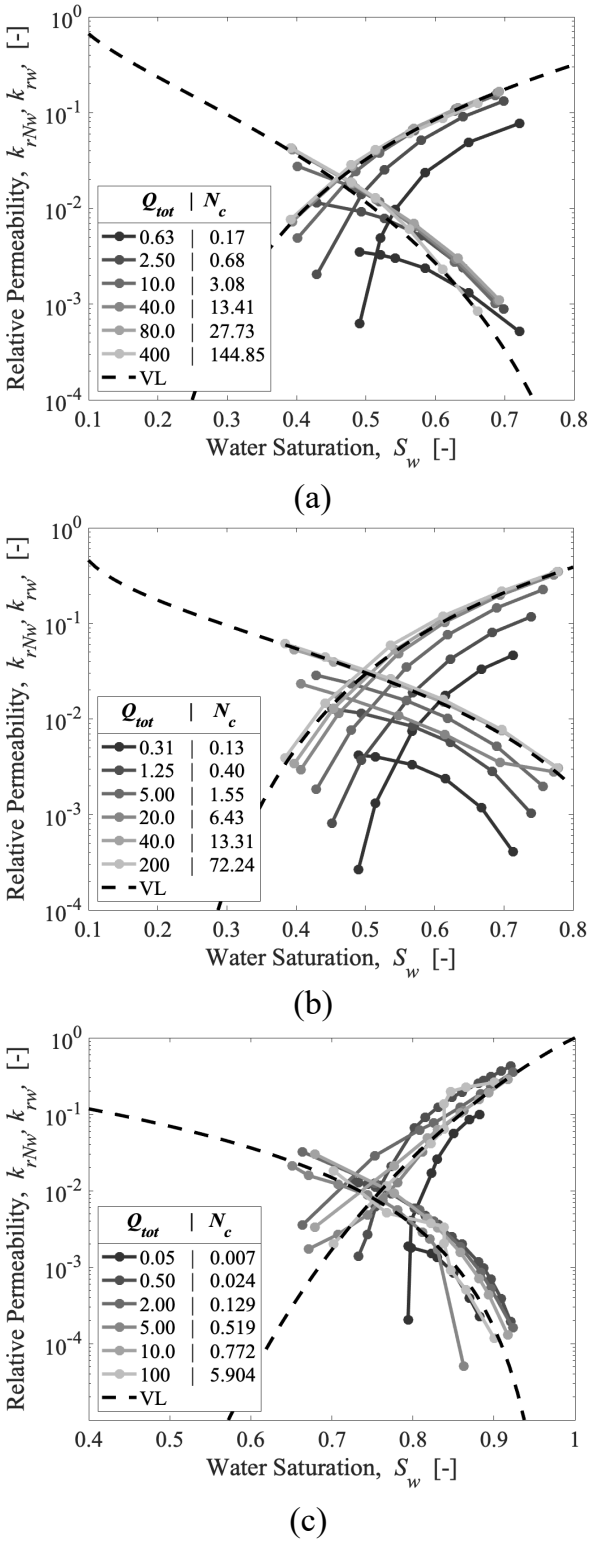

Figure 10. Equivalent relative permeabilities for the (a) Bentheimer sandstone, (b) Bunter sandstone and (c) Edwards Brown dolomite samples at varying flow rates. The tabulated data is presented in Tables S1-S3 in the supplementary material. VL stands for the viscous-limit relative permeability used as input to the simulations. The samples display contrasting rate dependency behaviour due to a range of heterogeneity characteristics.

4 Conclusions

We have applied the capillary heterogeneity characterisation workflow presented in Jackson et al. (2018) to five samples with varying degrees of heterogeneity. This allowed us to test the applicability of the workflow to carbonate samples with more complex heterogeneities. Additionally, a direct comparison of the rate dependency of relative permeability in various rock types could be made.

The workflow successfully characterised the heterogeneity in the two sandstones, where a good match in the experiment and simulation measurements was observed. In contrast to this, the successful application of the workflow to the carbonates was found to correlate with the size of the heterogeneous features. The Indiana limestone with small variations ($\approx$ 1 mm) in porosity, permeability and entry pressure led to the poorest match. The simulations could not match the saturation distribution, significantly underpredicted the gas relative permeability, and displayed near to no rate dependency. The Estaillades carbonate with larger features allowed for a match in the saturation distribution but an underprediction of the relative permeability, and an incorrect estimate of the impact of flow rate. The Edwards Brown with large heterogeneities that were clearly resolved by the imagery exhibited a good agreement in the relative permeabilities. This supports the hypothesis that the iterative calibration is dependent on the amount of controlling rock structure that can be resolved with the medical CT scanner.

To improve the predictions, we tested a fully empirical approach whereby the input viscous-limit relative permeabilities were used as fitting parameters. This approach improved the matches of saturation distributions at a target flow rate, but failed to result in predictive capability for the impact of flow rate on saturation distribution or relative permeability in all three carbonates. Notably, it led to the loss of predictive capability achieved with the initial iterative approach on the Edwards Brown. This finding emphasizes the importance of considering the physical mechanisms controlling the fluid behaviour in the creation of a numerical model of the rock cores. The purely numerical approach to matching the experiments led to a complete loss of the predictive capabilities of the models.

For the samples with a successful model, the Bentheimer sandstone, Bunter sandstone and Edwards Brown dolomite, we investigated the flow rate dependency of the upscaled relative permeability. We found that parallel layering as present in the Bentheimer sample allows for the phases to distribute optimally within the pore space, thereby raising the gas relative permeability. Perpendicular bedding in the Bunter sandstone hinders optimal flow of either phases, hence reducing the relative permeabilities. The isotropic heterogeneity in the Edwards Brown resulted in non-monotonic behaviour; initially relative permeability was increased, and subsequently decreased with increasing flow rate.

The work both underscores the importance of capturing small-scale heterogeneities in characterising subsurface fluid flows, as well as the challenges in doing so. Where imagery can sufficiently resolve heterogeneous features, the 3D history match procedure can be performed entirely within the realm of Darcy based reservoir simulators. For rocks where smaller scale features evidently place controls on the upscaled flow further techniques may be required such as the use of pore-network or other pore-scale flow models (Zahasky et al., 2020).

Acknowledgments

This work was funded by the Engineering and Physical Sciences Research Council (iCase studentship). We acknowledge PRORES and Computer Modelling Group (CMG) for providing access to SENDRA and IMEX, respectively. The experimental sandstone and carbonate data associated with this work may be obtained from the following data repositories: `http://www.bgs.ac.uk/ukccs/accessions/index.html#item107811` and `https://www2.bgs.ac.uk/ukccs/accessions/index.html#item133485`, respectively. The numerical models used in this work are currently being uploaded. We thank Sojwal Manoorkaar for providing the carbonate coreflood datasets and for useful discussions.

References

Al-Kharusi, A. S., & Blunt, M. J. (2008). Multiphase flow predictions from carbonate pore space images using extracted network models. *Water Resources Research*, *44*(6), 1–14. doi: 10.1029/2006WR005695

Al-Menhali, A. S., Menke, H. P., Blunt, M. J., & Krevor, S. C. (2016). Pore Scale Observations of Trapped CO_2 in Mixed-Wet Carbonate Rock: Applications to Storage in Oil Fields. *Environmental Science and Technology*, *50*(18), 10282–10290. doi: 10.1021/acs.est.6b03111

Aminu, M. D., Nabavi, S. A., Rochelle, C. A., & Manovic, V. (2017). A review of developments in carbon dioxide storage. *Applied Energy*, *208*(August), 1389–1419. doi: 10.1016/j.apenergy.2017.09.015

Arts, R., Eiken, O., Chadwick, A., Zweigel, P., van der Meer, L., & Zinszner, B. (2004). Monitoring of CO_2 injected at Sleipner using time-lapse seismic data. *Energy*, *29*(9-10), 1383–1392. doi: 10.1016/j.energy.2004.03.072

Benham, G., Bickle, M., & Neufeld, J. (2020). Upscaling multiphase flow through heterogeneous porous media. , 1–31.

Berg, S., Oedai, S., & Ott, H. (2013). Displacement and mass transfer between saturated and unsaturated CO_2-brine systems in sandstone. *International Journal of Greenhouse Gas Control*, *12*, 478–492. doi: 10.1016/j.ijggc.2011.04.005

Berg, S., Unsal, E., & Dijk, H. (2021). Non-uniqueness and uncertainty quantification of relative permeability measurements by inverse modelling. *Computers and Geotechnics*, *132*, 1–18. doi: 10.1016/j.compgeo.2020.103964

Birkholzer, J. T., Oldenburg, C. M., & Zhou, Q. (2015). CO_2 migration and pressure evolution in deep saline aquifers. *International Journal of Greenhouse Gas Control*, *40*, 203–220. doi: 10.1016/j.ijggc.2015.03.022

Brook, M., Shaw, K., Vincent, C., & Holloway, S. (2003). Gestco case study 2a-1: Storage Potential of the Bunter Sandstone in the UK sector of the Southern North Sea and the adjacent onshore area of Eastern England. *British Geological Survey Commissioned Report CR/03/154N*, 37.

Brooks, R. H., & Corey, A. T. (1964). Hydraulic Properties of Porous Media. *Colorado State University: Hydrology Papers*, *3*(1), 1–23.

Cantrell, D. L., & Hagerty, R. M. (1999). Microporosity in Arab Formation Carbonates, Saudi Arabia. *GeoArabia*, *4*(2), 129–154.

Chadwick, R. A., Noy, D., Arts, R., & Eiken, O. (2009). Latest time-lapse seismic data from Sleipner yield new insights into CO_2 plume development. *Energy Procedia*, *1*(1), 2103–2110. doi: 10.1016/j.egypro.2009.01.274

Chang, J., & Yortsos, Y. (1992). Effect of capillary heterogeneity on Buckley-Leverett displacement. *SPE Reservoir Engineering*, *7*, 285 – 293.

Chaouche, M., Rakotomalala, N., Salin, D., Xu, B., & Yortsos, Y. C. (1994). Capillary effects in drainage in heterogeneous porous media: continuum modelling, experiments and pore network simulations. *Chemical Engineering Science*, *49*(15), 2447–2466. doi: 10.1016/0009-2509(94)E0040-W

Chierici, G. L. (1984). Novel relations for drainage and imbibition relative permeabilities. *SPE Journal*, *24*, 275–276.

Corey, A., & Rathjens, C. (1956). Effect of Stratification on Relative Permeability. *Journal of Petroleum Technology*, *8*, 69 – 71. doi: 10.2118/744-g

Dawe, R. A., Wheat, M. R., & Bidner, M. S. (1992). Experimental investigation of capillary pressure effects on immiscible displacement in lensed and layered porous media. *Transport in Porous Media*, *7*, 83 – 101.

El-Maghraby, R. M. (2012). Measurements of CO_2 Trapping in Carbonate and Sandstone Rocks - Dissertation. *Imperial College London*.

Global CCS Institute. (2019). *Global status of ccs: 2019*.

Green, C. P., & Ennis-King, J. (2010). Effect of vertical heterogeneity on long-term migration of CO_2 in saline formations. *Transport in Porous Media*, *82*(1), 31–47.

Hamon, G., & Roy, C. (2000). Influence of Heterogeneity , Wettability and Coreflood Design on Relative Permeability Curves. *Paper SCA*(1978).

Heidsiek, M., Butscher, C., Blum, P., & Fischer, C. (2020). Small-scale diagenetic facies heterogeneity controls porosity and permeability pattern in reservoir sandstones. *Environmental Earth Sciences*, *79*(425), 1–14. doi: 10.1007/s12665-020-09168-z

Hosseini, S. A., Lashgari, H., Choi, J. W., Nicot, J. P., Lu, J., & Hovorka, S. D. (2013). Static and dynamic reservoir modeling for geological CO_2 sequestration at Cranfield, Mississippi, U.S.A. *International Journal of Greenhouse Gas Control*, *18*, 449–462. doi: 10.1016/j.ijggc.2012.11.009

Hosseinzadeh Hejazi, S. A., Shah, S., & Pini, R. (2019). Dynamic measurements of drainage capillary pressure curves in carbonate rocks. *Chemical Engineering Science*, *200*, 268–284. doi: 10.1016/j.ces.2019.02.002

Hove, A. O., Nilsen, V., & Jorgen, L. (1990). Visualization of Xanthan flood behavior in core samples by means of X-ray tomography. *SPE Reservoir Engineering*, *5*, 475 – 480.

Huang, Y., Ringrose, P. S., & Sorbie, K. S. (1995). Capillary trapping mechanisms in water-wet laminated rocks. *SPE Reservoir Engineering (Society of Petroleum Engineers)*, *10*(4), 287–292. doi: 10.2118/28942-PA

Huppler, J. D. (1970). Numerical investigation of the effects of core heterogeneities on waterflood relative permeabilities. *AIME Petroleum Transcations*, *10*, 381 – 392.

Jackson, S. J., Agada, S., Reynolds, C. A., & Krevor, S. (2018). Characterizing Drainage Multiphase Flow in Heterogeneous Sandstones. *Water Resources Research*, *54*(4), 3139–3161. doi: 10.1029/2017WR022282

Jackson, S. J., & Krevor, S. (2020). Small-Scale Capillary Heterogeneity Linked to Rapid Plume Migration During CO_2 Storage. *Geophysical Research Letters*, *47*(18). doi: 10.1029/2020GL088616

Jackson, S. J., Lin, Q., & Krevor, S. (2020). Representative Elementary Volumes, Hysteresis, and Heterogeneity in Multiphase Flow From the Pore to Continuum Scale. *Water Resources Research*, *56*(6), 1–24. doi: 10.1029/2019WR026396

Kong, X., Delshad, M., & Wheeler, M. F. (2015). History matching heterogeneous coreflood of CO_2/brine by use of compositional reservoir simulator and geostatistical approach. *SPE Journal*, *20*(2), 267–276. doi: 10.2118/163625-pa

Kortekaas, T. F. (1985). Water/Oil Displacement Characteristics in Crossbedded Reservoir Zones. *Society of Petroleum Engineers journal*, *25*(6), 917–926. doi: 10.2118/12112-PA

Krause, M., & Benson, S. M. (2015). Accurate determination of characteristic relative permeability curves. *Advances in Water Resources*, *83*, 376–388. doi: 10.1016/j.advwatres.2015.07.009

Krause, M., Krevor, S., & Benson, S. M. (2013). A Procedure for the Accurate Determination of Sub-Core Scale Permeability Distributions with Error Quantification. *Transport in Porous Media*, *98*(3), 565–588. doi:

10.1007/s11242-013-0161-y

Krause, M., Perrin, J. C., & Benson, S. M. (2011). Modeling permeability distributions in a sandstone core for history matching coreflood experiments. *SPE Journal*, *16*(4), 768–777. doi: 10.2118/126340-PA

Krevor, S., Blunt, M. J., Benson, S. M., Pentland, C. H., Reynolds, C., Al-Menhali, A., & Niu, B. (2015). Capillary trapping for geologic carbon dioxide storage - From pore scale physics to field scale implications. *International Journal of Greenhouse Gas Control*, *40*, 221–237. doi: 10.1016/j.ijggc.2015.04.006

Krevor, S., Blunt, M. J., Trusler, J., & De Simone, S. (2019). An introduction to subsurface CO_2 storage. In *Carbon capture and storage* (Vol. 26, pp. 238–295). Cambridge: Royal Society of Chemistry.

Krevor, S. C., Pini, R., Li, B., & Benson, S. M. (2011). Capillary heterogeneity trapping of CO_2 in a sandstone rock at reservoir conditions. *Geophysical Research Letters*, *38*(15), 1–5. doi: 10.1029/2011GL048239

Lai, P., Moulton, K., & Krevor, S. (2015). Pore-scale heterogeneity in the mineral distribution and reactive surface area of porous rocks. *Chemical Geology*, *411*, 260–273. doi: 10.1016/j.chemgeo.2015.07.010

Lu, J., Kordi, M., Hovorka, S. D., Meckel, T. A., & Christopher, C. A. (2013). Reservoir characterization and complications for trapping mechanisms at Cranfield CO_2 injection site. *International Journal of Greenhouse Gas Control*, *18*, 361–374. doi: 10.1016/j.ijggc.2012.10.007

Manoorkaar, S., Jackson, S. J., & Krevor, S. (2021). Observations of the impacts of 0.001m - 0.1m heterogeneities on relative permeability and trapping in carbonate rocks. *Water Resources Research*.

Ni, H., Boon, M., Garing, C., & Benson, S. M. (2019). Predicting CO_2 residual trapping ability based on experimental petrophysical properties for different sandstone types. *International Journal of Greenhouse Gas Control*, *86*(May), 158–176. doi: 10.1016/j.ijggc.2019.04.024

Onoja, M. U., & Shariatipour, S. M. (2019). Assessing the impact of relative permeability and capillary heterogeneity on Darcy flow modelling of CO_2 storage in Utsira Formation. *Greenhouse Gases: Science and Technology*, *9*(6), 1221–1246. doi: 10.1002/ghg.1932

Peksa, A. E., Wolf, K. H. A., & Zitha, P. L. (2015). Bentheimer sandstone revisited for experimental purposes. *Marine and Petroleum Geology*, *67*, 701–719. doi: 10.1016/j.marpetgeo.2015.06.001

Perrin, J. C., & Benson, S. (2010). An experimental study on the influence of subcore scale heterogeneities on CO_2 distribution in reservoir rocks. *Transport in Porous Media*, *82*(1), 93–109. doi: 10.1007/s11242-009-9426-x

Petrovic, A., Aigner, T., & Pontiggia, M. (2018). Facies Heterogeneities in a Ramp Carbonate Reservoir Analogue: a New High-Resolution Approach for 3D Facies Modelling. *Journal of Petroleum Geology*, *41*(2), 155–174. doi: 10.1111/jpg.12698

Pini, R., & Benson, S. M. (2013a). Characterization and scaling of mesoscale heterogeneities in sandstones. *Geophysical Research Letters*, *40*(15), 3903–3908. doi: 10.1002/grl.50756

Pini, R., & Benson, S. M. (2013b). Simultaneous determination of capillary pressure and relative permeability curves from core-flooding experiments with various fluid pairs. *Water Resources Research*, *49*(6), 3516–3530. doi: 10.1002/wrcr.20274

Pini, R., Krevor, S. C., & Benson, S. M. (2012). Capillary pressure and heterogeneity for the CO 2/water system in sandstone rocks at reservoir conditions. *Advances in Water Resources*, *38*, 48–59.

Prodanović, M., Mehmani, A., & Sheppard, A. P. (2015). Imaged-based multiscale network modelling of microporosity in carbonates. *Geological Society Special Publication*, *406*(1), 95–113. doi: 10.1144/SP406.9

Rabinovich, A., Li, B., & J., D. L. (2016). Analytical approximations for effective relative permeability in the capillary limit Avinoam. *Journal of the American Water Resources Association*, *5*(3), 2–2. doi: 10.1111/j.1752-1688.1969.tb04897.x

Rebelle, M., Umbhauer, F., & Poli, E. (2009). IPTC 13120 Pore to Grid Carbonate Rock-Typing Geological Facies Petro-Geologic Groups (PGG) Petrophysical Groups (PG) Dynamic Rock-Types (DRT) Core to log. *Petrophysics*(December), 7–9.

Reynolds, C. A., Blunt, M. J., & Krevor, S. (2018). Multiphase Flow Characteristics of Heterogeneous Rocks From CO_2 Storage Reservoirs in the United Kingdom. *Water Resources Research*, *54*(2), 729–745. doi: 10.1002/2017WR021651

Reynolds, C. A., & Krevor, S. (2015). Characterizing flow behavior for gas injection: Relative permeability of CO_2-brine and N_2-water in heterogeneous rocks. *Water Resources Research*. doi: 10.1002/2015WR017200.A

Ringrose, P. S., Sorbie, K. S., Corbett, P. W., & Jensen, J. L. (1993). Immiscible flow behaviour in laminated and cross-bedded sandstones. *Journal of Petroleum Science and Engineering*, *9*(2), 103–124. doi: 10.1016/0920-4105(93)90071-L

Sarwaruddin, M., Skauge, A., & Torsaeter, O. (2001). Modeling of Capillary Pressure for Heterogeneous Reservoirs By a Modified J-Function. (September 2014).

Sayers, C. M. (2008). The elastic properties of carbonates. *The Leading Edge*, *27*(8), 1020–1024. doi: 10.7588/worllitetoda.89.5.0034

Virnovsky, G. A., Friis, H. A., & Lohne, A. (2004). A Steady-State Upscaling Approach for Immiscible Two-Phase Flow. *Transport in Porous Media*, *54*(2), 167–192. doi: 10.1023/A:1026363132351

Withjack, E. M. (1988). Computed tomography for rock-property determination and fluid-flow visualization. *SPE Formation Evaluation*, *3*(4), 696–704. doi: 10.2118/16951-pa

Woods, A. W., & Farcas, A. (2009). Capillary entry pressure and the leakage of gravity currents through a sloping layered permeable rock. *Journal of Fluid Mechanics*, *618*, 361–379.

Worden, R. H., Armitage, P. J., Butcher, A. R., Churchill, J. M., Csoma, A. E., Hollis, C., ... Omma, J. E. (2018). Petroleum reservoir quality prediction: overview and contrasting approaches from sandstone and carbonate communities. *Geological Society, London, Special Publications*, *435*(1), 1–31. doi: 10.1144/SP435.21

Yamamoto, H. (2009). Numerical investigation for the impact of CO_2 geologic sequestration on regional groundwater flow. *International Journal of Greenhouse Gas Control*, 35–43.

Zahasky, C., & Benson, S. M. (2018). Micro-positron emission tomography for measuring sub-core scale single and multiphase transport parameters in porous media. *Advances in Water Resources*, *115*, 1–16. doi: 10.1016/j.advwatres.2018.03.002

Zahasky, C., Jackson, S. J., Lin, Q., & Krevor, S. (2020). Pore Network Model Predictions of Darcy-Scale Multiphase Flow Heterogeneity Validated by Experiments. *Water Resources Research*, *56*(6), 1–16. doi: 10.1029/2019WR026708