

# On the Validation of a Compositional Model for the Simulation of CO<sub>2</sub> Injection into Saline Aquifers

Morteza Akbarabadi<sup>1</sup> · Marcio Borges<sup>2</sup> · Ahmad Jan<sup>3</sup> · Felipe Pereira<sup>4</sup> · Mohammad Piri<sup>1</sup>

Received: 4 October 2015 / Accepted: 8 May 2017 / Published online: 1 June 2017  
© Springer Science+Business Media Dordrecht 2017

**Abstract** In this work we apply a recently proposed Bayesian Markov chain Monte Carlo framework (Akbarabadi et al. in *Comput Geosci* 19(6):1231–1250, 2015) to quantify uncertainty in the three-dimensional permeability field of a rock core. This process establishes the credibility of a compositional two-phase flow model to describe the displacement of brine by CO<sub>2</sub> and CO<sub>2</sub> storage in saline aquifers. We investigate the predictive capabilities of the compositional model in the context of an unsteady-state CO<sub>2</sub>-brine drainage experiment at the laboratory scale, performed at field-scale aquifer conditions. We employ forward models consisting of a system of discretized partial differential equations along with relative permeability curves obtained by a curve fitting of experimental measurements. We consider a forward model to be validated when: (1) numerical simulations reveal that the Bayesian framework has accurately characterized the core's permeability and (2) Monte Carlo predictions show excellent agreement between measured and simulated data. A large set of numerical studies with an accurate compositional simulator shows that forward models have been successfully validated. For such models, our numerical results show that we are able to capture all the dominant features and general trends of the CO<sub>2</sub> saturation fields observed in the core. Our study is consistent with the design and findings of real experiments. Fluid properties, relative permeability data, measured porosity field, physical dimensions, and thermodynamic conditions are the same as those reported in Akbarabadi and Piri (*Adv Water Resour* 52:190–206, 2013). However, the measured saturation data are from flow experiments different from those reported in Akbarabadi and Piri (2013), and will be presented here.

---

✉ Felipe Pereira  
luisfelipe.pereira@utdallas.edu

<sup>1</sup> Department of Chemical and Petroleum Engineering, University of Wyoming, Laramie, WY 82071, USA

<sup>2</sup> National Laboratory for Scientific Computing, Petropolis, RJ 25651, Brazil

<sup>3</sup> Department of Mathematics, University of Wyoming, Laramie, WY 82071, USA

<sup>4</sup> Mathematical Sciences Department, University of Texas at Dallas, Richardson, TX 75080, USA

**Keywords** Validation · Core characterization · Bayesian MCMC · Compositional model · Drainage · CO<sub>2</sub> injection

## 1 Introduction

CO<sub>2</sub> is a greenhouse gas of considerable concern in connection with global warming. As a method to reduce the emission rate of anthropogenic CO<sub>2</sub> into the atmosphere, the storage of CO<sub>2</sub> into geologic formations has attracted a lot of attention, because these formations offer a significant potential for long-term storage of CO<sub>2</sub>. This paper examines a method for validating mathematical models of the fluid displacement processes involved in this technology.

The purpose of performing CO<sub>2</sub>-brine drainage experiments at the core scale is to help improve our understanding of fluid flow and displacement of brine by CO<sub>2</sub> in saline aquifers. Recent studies show that rock heterogeneity has a strong influence on the saturation patterns of brine displaced by CO<sub>2</sub> in cores. Even for the Berea sandstone, which is considered to be a relatively homogeneous rock (Krause et al. 2011), laboratory experiments show that for a relative low variability of the porosity, the corresponding CO<sub>2</sub> saturation profile is highly heterogeneous (Akbarabadi and Piri 2013; Shi et al. 2009; Benson et al. 2006). These experiments provide insight into the role of heterogeneity in determining distribution of the CO<sub>2</sub> saturation.

The agreement between core-scale numerical simulations and experimentally measured fluid properties, such as saturation, has been and will continue to be a challenge because of uncertainty in the heterogeneities in rock and fluid properties. Accurate numerical simulations of core-flooding experiments can be used to help in quantifying the underlying uncertainty associated with geological properties such as permeability. The main challenge in validating mathematical models against core-flooding experiments is to find the 3-dimensional (3D) permeability field of a core. One cannot measure the 3D permeability field in laboratory experiments directly, in contrast to other functions such as porosity, relative permeability curves, etc. Instead, one must start with statistics of the permeability field, then construct realizations that are consistent with the observed behavior of fluids in the rock. Finding accurate 3D permeability field is essential for a successful model validation that establishes the credibility of mathematical models.

In the work presented here, we apply a recently proposed Bayesian Markov chain Monte Carlo (MCMC) framework (Akbarabadi et al. 2015) to quantify uncertainty in the 3D permeability field of a core in the context of an unsteady-state CO<sub>2</sub>-brine drainage experiment. However, a more challenging situation is reported in this work, because the modeling and numerical errors (Glimm and Sharp 1999) were not present in the work authors reported in Akbarabadi et al. (2015) (synthetic computational experiments using virtual cores). Additional uncertainties related to the models for flow, relative permeability curves, fluid properties, etc., have great impact over prediction capacity. The experiment is conducted at the laboratory scale, at aquifer conditions. The Bayesian framework incorporates the measured CO<sub>2</sub> saturation data at transverse slices along the core, and consists of two steps: the characterization and predictions. The core characterization step is based on selecting a suitable family of 3D permeability fields such that solutions of a forward model, consisting of a system of discretized partial differential equations along with relative permeability curves, match the measured data within some error. In this work, the system of partial differential equations is a compositional two-phase flow model accounting for mass transfer between phases, and sets of relative permeability curves are constructed from the laboratory measurements by curve fitting procedures. The predictive capability of the forward model is

quantified through Monte Carlo simulations using the selected permeability fields from the posterior distribution.

We consider a forward model to be validated when: (1) numerical simulations reveal that the Bayesian framework has accurately characterized the core and (2) Monte Carlo predictions show excellent agreement between measured and simulated data. A large set of numerical studies confirms that our forward models have been successfully validated. However, we also found that forward models may not be adequate to describe the experimental results considered here. Such models fail to properly characterize the core and/or produce poor predictions of future fluid flow. For the validated models, our numerical results capture all of the dominant features and general trends of distribution of the measured CO<sub>2</sub> saturation in the core. Our numerical studies also show that a change in the relative permeability curves may also affect the average permeability of the proposals from the posterior distribution.

The pertinent literature to the problem at hand has mostly emphasized the use of indirect methods to estimate the permeability field of rocks. Most of the earlier work for permeability estimation has focused on expressing it in terms of other measurable rock properties. In particular, these indirect methods incorporate porosity, grain size, surface area, and pore dimension to determine the permeability field, see [Nelson \(1994\)](#). Many authors have studied different techniques to validate mathematical models to describe CO<sub>2</sub>-brine drainage in core floods. The authors of [Krause et al. \(2011\)](#) performed numerical simulations using various porosity-based permeability models to estimate the permeability distribution. Simulated results presented in [Krause et al. \(2011\)](#) indicate that a better model is needed to obtain a more accurate estimate of the permeability field. The authors of [Kong et al. \(2014\)](#) reported a geostatistical approach to estimate permeability using the porosity field and the capillary pressure function. Good agreement between numerical results and the measured data [reported in [Krause et al. \(2011\)](#)] was reported for a steady-state CO<sub>2</sub>-brine drainage experiment. In order to accurately characterize the core, a new model was used for the local capillary pressure to estimate sub-core-scale permeability distribution in [Krause et al. \(2013\)](#). The numerical results agree well with the measured data. We remark that our framework does not assume a dependence of permeability on other rock properties. However, such dependence in principle could be incorporated into the Bayesian framework. In addition, we aim at validating a mathematical model to describe an unsteady-state CO<sub>2</sub>-brine drainage experiment, in contrast to the published studies for steady-state experiments. For a detailed review of published research on model validation the reader is referred to [Kong et al. \(2014\)](#).

The paper proceeds as follows. Section 2 briefly describes the experimental procedure and illustrates the measured data. Section 3 is dedicated to the mathematical model. Section 4 reviews the Bayesian framework for model validation described in [Akbarabadi et al. \(2015\)](#). Section 5 presents validation results and discussion. A summary and concluding remarks appear in Sect. 6.

## 2 Laboratory Experiments

### 2.1 Experimental Methodology

This subsection presents a brief review of the design and setup of the CO<sub>2</sub>-brine core-flooding experiments. The experimental methodology that we follow in this work is exactly the same as reported in [Akbarabadi and Piri \(2013\)](#). However, as stated earlier the measured saturation data are from flow experiments different from those reported in [Akbarabadi and](#)

**Table 1** Total pore volume injected for each scan

| Scan no. | Pore volume injected |
|----------|----------------------|
| 1        | 0.6132               |
| 2        | 0.7622               |
| 3        | 0.9110               |
| 4        | 1.0597               |
| 5        | 1.2283               |
| 6        | 1.5307               |
| 7        | 1.8227               |

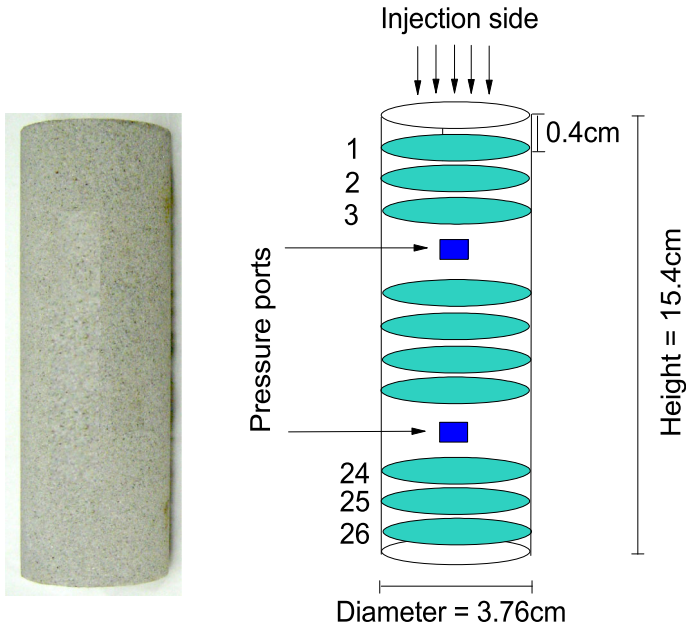
Piri (2013). These experiments involve an unsteady-state CO<sub>2</sub>-brine drainage conducted in a Berea sandstone at the laboratory scale. In the experiment a rock core of height 15.4 cm and diameter 3.76 cm was placed in a Hassler-type core holder, held vertically, and sealed in a sleeve. To prevent CO<sub>2</sub> from leaking through the sleeve, both the core and sleeve were wrapped with several layers of aluminum foil and Teflon tape. The core holder was then placed in the gantry of a medical CT scanner after wrapping with heating tapes and highly efficient insulation. The core was then heated to the temperature of 55 °C, and pressurized to 11 MPa. After that the core was fully saturated with brine and then flooded from the top with 100% equilibrated supercritical CO<sub>2</sub> at a constant injection rate of 0.15 cm<sup>3</sup> (cc)/min. A medical CT scanner was used to scan the core at several times to obtain the saturation of CO<sub>2</sub> in transverse slices of the core. The CT scanner provides a resolution of about 250 μm per slice with a spacing of about 4 mm, i.e., slices are 4 mm apart from each other. One complete scan (the term scan refers to the measurements made with a CT scanner at a particular time) generates data at 37 slices within about 5 min. Table 1 presents the number of scans and total pore volume injected for each scan.

The Berea sandstone core used in the experiment and the corresponding spatial domain for the governing equations are shown in Fig. 1. The inlet is subject to a Neumann boundary condition (a constant injection rate), and the outlet obeys a Dirichlet boundary condition (constant pressure). We impose no-flow conditions on all other segments of the boundary. In the experiment, two pressure ports were used along the length of the core to monitor the pressure. These ports affect the measurements, and as a consequence the CT scanner does not provide physically meaningful CO<sub>2</sub> saturation and porosity data close to these pressure ports. Therefore, we ignore data over this length and those of a few slices on each side of the two pressure ports. Moreover, there are some slices where the scanner-determined CO<sub>2</sub> saturation is negative with magnitude 10<sup>-3</sup>–10<sup>-1</sup>. We do not ignore such slices; instead, we round the negative saturation values to zero. In most cases, such slices are close to the outlet. Table 2 shows physical positions and slice numbers where we used the scanner data. These are the slices used to compare against numerical simulations. Slices associated with inaccurate measurements are not numbered but highlighted in italic.

## 2.2 Experimental Data

### 2.2.1 Measured Porosity and CO<sub>2</sub> Saturation Data

We present the measured porosity values and the CO<sub>2</sub> saturation data at some selected slices of the core to highlight the importance of heterogeneities in the spatial distribution of the CO<sub>2</sub> saturation. The Berea sandstone is considered to be a homogeneous rock in the sense

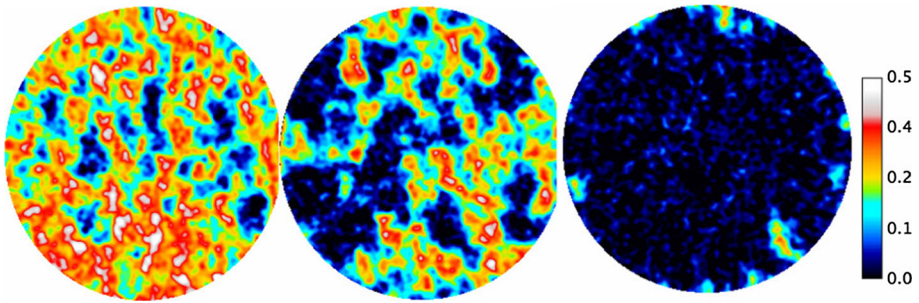


**Fig. 1** (Left) The Berea sandstone used in the experiment; (Right) Core description: Pressure ports (in blue) and slices for monitoring CO<sub>2</sub> saturation (in light green). Injection of supercritical CO<sub>2</sub> in the direction of gravity

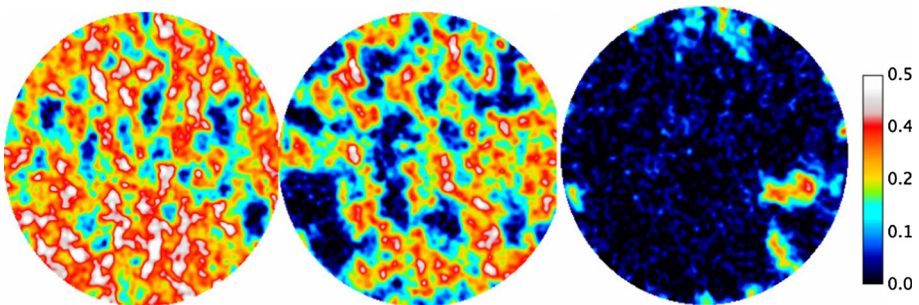
**Table 2** Physical position and slices numbers

| Slice no. | Physical position (cm) | Slice no. | Physical position (cm) | Slice no. | Physical position (cm) |
|-----------|------------------------|-----------|------------------------|-----------|------------------------|
| 1         | 15.0                   | 9         | 9.80                   | 22        | 4.60                   |
| 2         | 14.6                   | 10        | 9.40                   | 23        | 4.20                   |
| 3         | 14.2                   | 11        | 9.00                   | –         | 3.80                   |
| –         | 13.8                   | 12        | 8.60                   | –         | 3.40                   |
| –         | 13.4                   | 13        | 8.20                   | –         | 3.00                   |
| –         | 13.0                   | 14        | 7.80                   | –         | 2.60                   |
| –         | 12.6                   | 15        | 7.40                   | –         | 2.20                   |
| –         | 12.2                   | 16        | 7.00                   | –         | 1.80                   |
| 4         | 11.8                   | 17        | 6.60                   | 24        | 1.40                   |
| 5         | 11.4                   | 18        | 6.20                   | 25        | 1.00                   |
| 6         | 11.0                   | 19        | 5.80                   | 26        | 0.60                   |
| 7         | 10.6                   | 20        | 5.40                   |           |                        |
| 8         | 10.2                   | 21        | 5.00                   |           |                        |

that there are no apparent bedding layers in its structure. However, experiments show that the spatial distribution of the CO<sub>2</sub> saturation is very heterogenous even for this type of rock; see Benson et al. (2006), Shi et al. (2009), Perrin and Benson (2010), Shi et al. (2011), Akbarabadi and Piri (2013).



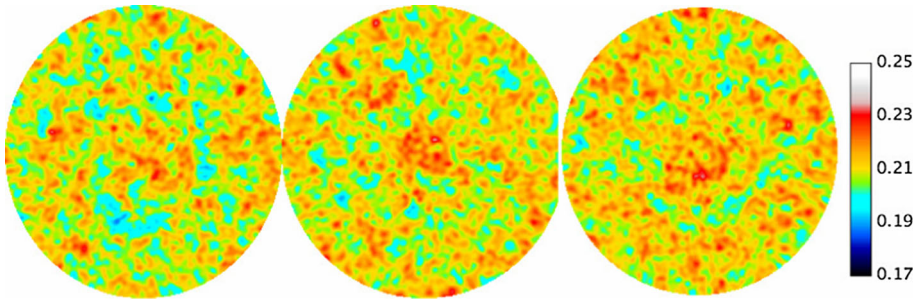
**Fig. 2** Measured  $\text{CO}_2$  saturation at locations close to the inlet boundary (*left*: slice # 3), in the middle part of the core (*middle*: slice # 13), and close to the outlet boundary (*right*: slice # 23) of scan 1. These data will be used for core characterization



**Fig. 3** Measured  $\text{CO}_2$  saturation at locations close to the inlet boundary (*left*: slice # 3), from the middle of the core (*middle*: slice # 13), and close to the outlet boundary (*right*: slice # 23) of scan 7. Data used to assess the accuracy of the predictive simulations

In Fig. 2, we show the measured  $\text{CO}_2$  saturation at slices # 3, 13, and 23, of the core of scan 1. The measured  $\text{CO}_2$  saturation field is highly heterogeneous. Along the length of the core, in the direction of injection, there is relatively less  $\text{CO}_2$ , especially at slices close to the bottom, and the  $\text{CO}_2$  is leaving the core in a channelized pattern. In the framework employed here, we use the measured data of scan 1 to characterize the core and later scans to assess the validity of the characterization. For example, Fig. 3 shows the measured data of scan 7 at slices # 3, 13, 23, which we use to assess the capability of a forward model to predict future fluid flow. In case of successful model validation, the numerically simulated results will match these data from scan 7, within an uncertainty, along with data from all other previous scans, which are not shown here. Figure 4 illustrates measured porosity data at slices # 3, 13, and 23. These data vary between 17 and 25%, with an average value of 21.2%. Note from Table 2 that slice # 3 is close to the inlet boundary, slice # 13 is in the middle part of the core, and slice # 23 is close to the outlet boundary of the core.

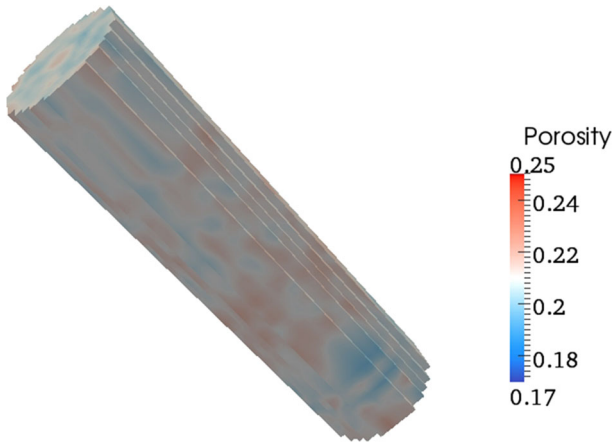
The numerical results presented here are consistent with a real experiment in that, for the simulations, we use the porosity field, fluid properties, relative permeability curves, injection rate, total injected pore volume for each scan, outlet boundary pressure, temperature, and physical dimensions that are reported in Akbarabadi and Piri (2013), see Tables 1, 2 and 3.



**Fig. 4** Measured porosity at locations close to the inlet boundary (left: slice # 3), from the middle of the core (middle: slice # 13), and close to the outlet boundary (right: slice # 23)

**Table 3** Summary of core and fluid properties (Akbarabadi and Piri 2013)

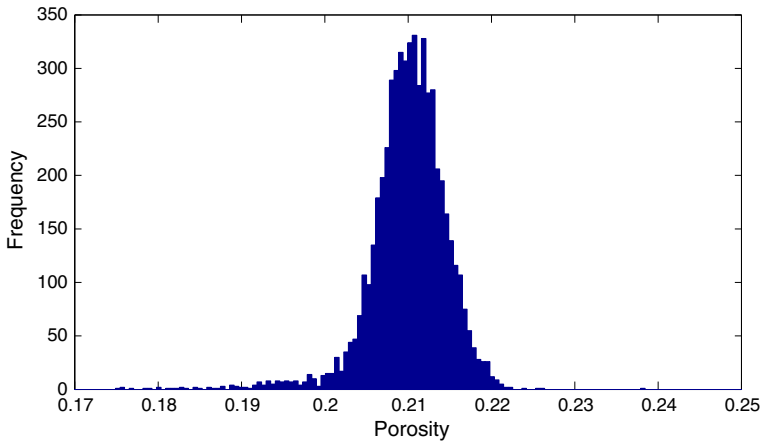
|                |         |                           |                        |
|----------------|---------|---------------------------|------------------------|
| Mean porosity  | 21.2%   | Injection rate            | 0.15 cc/min            |
| Length         | 15.4 cm | Brine density             | 1123 kg/m <sup>3</sup> |
| Diameter       | 3.76 cm | CO <sub>2</sub> density   | 393 kg/m <sup>3</sup>  |
| Pressure ports | 2       | Brine viscosity           | 9.1E-4 Pa s            |
| Back pressure  | 11 MPa  | CO <sub>2</sub> viscosity | 4.4E-5 Pa s            |
| Temperature    | 55 °C   | Pore volume               | 36.3 cc                |



**Fig. 5** An illustration of the measured porosity: three-dimensional profile with 16 × 16 × 64 grid elements excluding the dead (or inactive) cells

### 2.2.2 Construction of a 3D-Porosity Field for Simulations

In the simulations presented here, the porosity field is considered to be deterministic and is constructed from the measured porosity data at slices. Figure 5 shows the three-dimensional image of the porosity field used in simulations in a 16 × 16 × 64 grid. (We discuss the choice of grid size in Sect. 3.4.2.) For the cylindrical geometry, some of the cells are inactive and are not shown in Fig. 5. We linearly interpolate porosity values along the length of the core at the locations where porosity is not known from measurements. Figure 6 illustrates a histogram



**Fig. 6** Measured porosity histogram

of the measured porosity data exhibiting an approximately Gaussian distribution with an average of 0.212.

### 2.2.3 The Prior Distribution for the Permeability Field

In the Bayesian framework, we must provide the prior distribution of permeability for the characterization step. We assume that, although statistically independent, the porosity, and permeability fields share the same spatial structure. Thus, we first estimate the covariance of the measured porosity data and then adopt a best fit of these data as the covariance function for the (Gaussian) prior distribution of the permeability field. We find that the covariance function denoted by  $R(\mathbf{x}_1, \mathbf{x}_2)$  (see Sect. 4.1.3) with correlation length 0.1 cm in each direction matches the form of the covariance of the porosity field in  $x$ - and  $y$ -directions well (up to about 0.2 cm) as shown in Fig. 7 (for distances greater than 0.2 cm, we believe that we do not have enough data to estimate the correlation because the data show oscillations). In the  $z$ -direction, the smaller measured distance is 0.4 cm (the distance between two successive slices), and therefore, it is impossible to find correlation lengths smaller than this value. Thus, assuming isotropy of the rock, we consider the same correlation length (0.1 cm) in the  $z$ -direction. This finding motivates the use of  $R(\mathbf{x}_1, \mathbf{x}_2)$  as the covariance function for the prior distribution of the permeability field, assumed to be log-normal distributed and entirely characterized by its two-point covariance function (details in Sect. 4.1.3). In Fig. 7,  $C_Y$  represents covariance of the porosity data and  $r$  denotes the distance. The coefficient of determination is 1 which indicates  $R(\mathbf{x}_1, \mathbf{x}_2)$  is the best fit to the porosity covariance (using data up to the distance of 0.2 cm.) We remark that to the best of our knowledge the use of porosity correlation properties to define the prior for absolute permeability modeling has not appeared in the literature, and is used here for the first time.

### 2.2.4 Representative Elementary Volume (REV)

We use spatial mesh sizes for continuum-scale simulations that are larger than or comparable to the REV size. The size of the porosity-based REV for the core is computed using images with resolution  $2.34 \mu\text{m}$ . Porosity as a function of domain volume is illustrated in Fig. 8. This



**Fig. 7** An illustration of the covariance form of porosity and the exponential covariance function with correlation length 0.1 cm in each direction. **a**, **b**, and **c** correspond to  $x$ -,  $y$ -, and  $z$ -direction, respectively, and  $R^2$  is the coefficient of determination

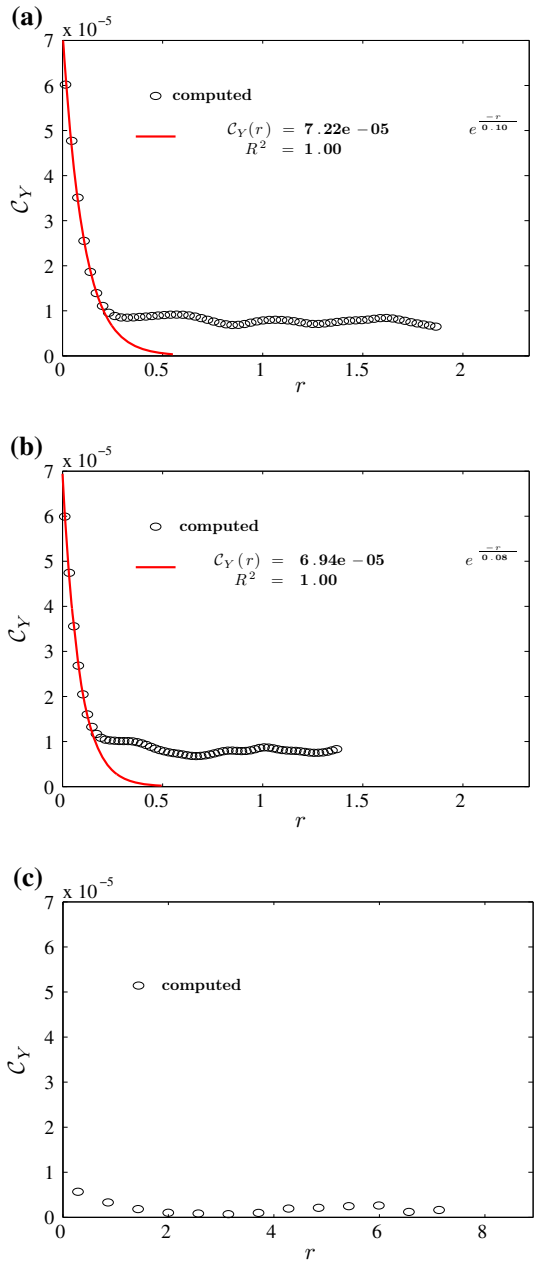
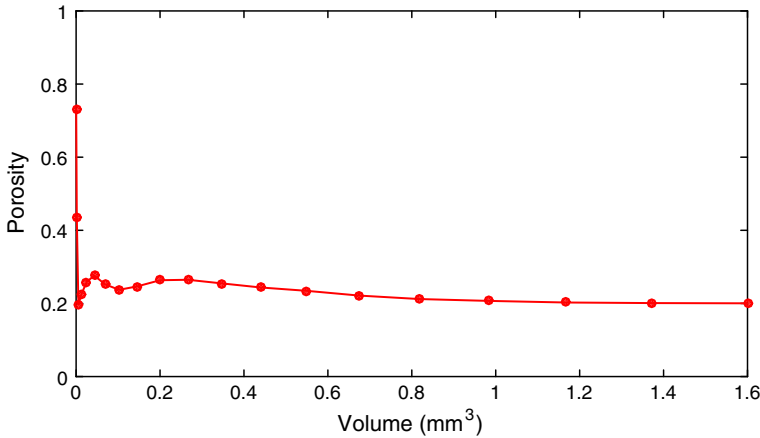


figure shows that oscillations in the porosity are not significant at averaging volumes larger than approximately  $0.78 \times 0.78 \times 0.78 \text{ mm}^3$ . This volume corresponds to  $335 \times 335 \times 335$  voxels. This observed size of the porosity REV is consistent with the data published in the literature; see [Mostaghimi et al. \(2013\)](#) and [Ovaysi and Piri \(2010\)](#).



**Fig. 8** (Representative elementary volume) An illustration of porosity as a function of volume (mm<sup>3</sup>) for images with resolution 2.34  $\mu\text{m}$

### 3 The Forward Model

#### 3.1 Governing Equations

We use a two-phase compositional model to describe the compressible fluid flow of supercritical CO<sub>2</sub> in saline formations. The two phases are referred to as aqueous (aq) and non-aqueous (naq) phases. There are three components distributed among the phases: CO<sub>2</sub> (*c*), water (*w*), and salt (*s*). The non-aqueous phase is the CO<sub>2</sub>-rich phase. For comprehensive reviews of compositional models, refer to [Trangenstein and Bell \(1989\)](#), [Qin \(1995\)](#), [Chen and Zhang \(2008\)](#), [Akbarabadi et al. \(2015\)](#), among many others. Here, we provide a brief summary for a quick reference.

We write a mass balance equation for each component in terms of the total moles per pore volume ([Trangenstein and Bell 1989](#); [Qin 1995](#)):

$$\phi \frac{\partial m^i}{\partial t} + \nabla \cdot \left( \frac{m_{\text{aq}}^i}{S_{\text{aq}}} \mathbf{v}_{\text{aq}} + \frac{m_{\text{naq}}^i}{S_{\text{naq}}} \mathbf{v}_{\text{naq}} \right) = 0, \quad i = c, w, s. \quad (1)$$

Here  $\mathbf{v}_{\text{aq}}$ ,  $\mathbf{v}_{\text{naq}}$ ,  $S_{\text{aq}}$ , and  $S_{\text{naq}}$  denote the velocities and saturations of the aqueous and non-aqueous phases, respectively. The variables  $m_{\text{aq}}^c$ ,  $m_{\text{aq}}^w$ ,  $m_{\text{aq}}^s$ ,  $m_{\text{naq}}^c$ ,  $m_{\text{naq}}^w$ , and  $m_{\text{naq}}^s$  are total moles per pore volume of each component in each phase. Note that there is no salt in the non-aqueous phase, that is,  $m_{\text{naq}}^s = 0$ . The superscripts *c*, *w*, *s* refer to the components CO<sub>2</sub>, water, and salt, while the subscripts aq and naq denote the aqueous and non-aqueous (CO<sub>2</sub>-rich) phases, respectively.

By mass conservation, the following relations hold:

$$m^i = m_{\text{aq}}^i + m_{\text{naq}}^i, \quad i = c, w, s. \quad (2)$$

The phase velocities  $\mathbf{v}_{\text{aq}}$  and  $\mathbf{v}_{\text{naq}}$  are functions of the corresponding phase saturation and pressure, given by Darcy's law for multiphase flow ([Bear 1979](#)):

$$\mathbf{v}_{\alpha} = -k\lambda_{\alpha}(S_{\text{aq}})(\nabla p_{\alpha} - \rho_{\alpha}\mathbf{g}\nabla z), \quad \alpha = \text{aq}, \text{naq}, \quad (3)$$

where  $\rho_\alpha$ ,  $\alpha = \text{aq, naq}$ , denotes phase mass density,  $k$  is the rock permeability,  $g$  is the gravity acceleration, and  $z$  is the depth. The functions  $\lambda_{\text{aq}}$  and  $\lambda_{\text{naq}}$  are the phase mobilities defined by

$$\lambda_\alpha(S_{\text{aq}}) = \frac{k_{r\alpha}}{\mu_\alpha} \quad \alpha = \text{aq, naq}, \quad (4)$$

where  $k_{r\alpha}$  and  $\mu_\alpha$ ,  $\alpha = \text{aq, naq}$ , denote the relative permeability and phase viscosities, respectively. We assume that the pore volume of the rock is fully filled with fluid, resulting in the constraint:

$$S_{\text{aq}} + S_{\text{naq}} = 1. \quad (5)$$

The compositional pressure equation used in this work, derived in [Trangenstein and Bell \(1989\)](#) and [Qin \(1995\)](#) using the volume balanced method, is given by:

$$\beta_T \frac{\partial p}{\partial t} + \text{div}(\mathbf{v}_t) = \phi \frac{S(t_n) - 1}{\Delta t}. \quad (6)$$

Here  $\mathbf{v}_t = \mathbf{v}_{\text{naq}} + \mathbf{v}_{\text{aq}}$  is the total Darcy velocity;  $S(t_n) = S_{\text{aq}} + S_{\text{naq}}$  is the total computed saturation,  $\phi$  is the porosity, and  $\Delta t$  is the time-step of the pressure equation. The parameter  $\beta_T$  is the total fluid compressibility, given by

$$\beta_T(p, \bar{\mathbf{m}}) = -\phi \left( \frac{\partial S_{\text{aq}}}{\partial p} + \frac{\partial S_{\text{naq}}}{\partial p} \right), \quad (7)$$

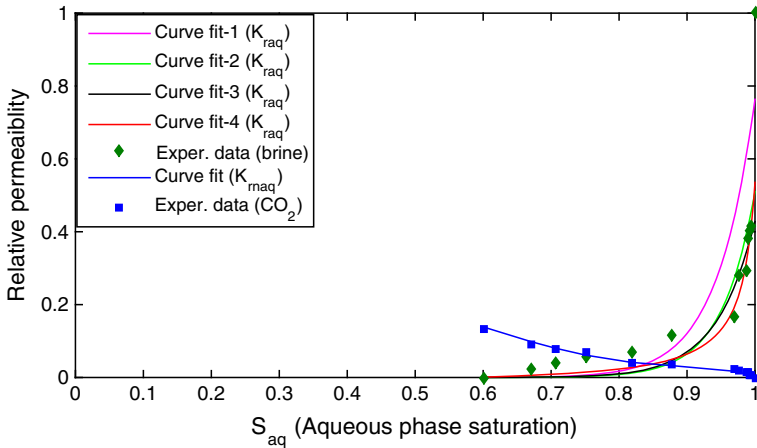
where  $\bar{\mathbf{m}} = \langle m^c, m^w, m^s \rangle$  is a vector of the total moles per pore volume of each component in the fluid. The pressure equation corresponds to a total volume balance. In computations, Eq. (5) may not hold exactly at a current time level  $t_n$ . The term on the right side of Eq. (7) serves to correct a possible volume discrepancy error. The phase pressures are related by capillary pressure function. For the numerical discretization strategy employed in our simulator, we refer to [Akbarabadi et al. \(2015\)](#).

### 3.2 Thermodynamics

For each grid cell in the spatial discretization, given the total number of moles of each component in the mixture, the pressure, and the temperature, the simulator computes the distribution of each component in each phase at the equilibrium state. This characterization of the fluid-phase equilibrium is known as a flash calculation. The flash calculation minimizes the total Gibbs free energy of the fluid mixture in each grid cell. The algorithm used in the minimization problem is given in [Akbarabadi et al. \(2015\)](#), and a detailed description of the flash calculation is given in [Leal \(2010\)](#).

### 3.3 Relative Permeabilities

The relative permeabilities are functions that determine how two fluids influence the motion of each other. We construct differentiable relative permeability curves that fit experimentally measured relative permeability data. The software used in this curve fitting estimates the fitting parameters using the least-squares method. Figure 9 shows the measured relative permeability data and fitted curves considered in this work. The analytical parameterizations of the relative permeability curves are given in ‘‘Appendix A’’. The mapping between discrete relative permeability data and differentiable curves is not unique, and the choice of curve has some effect on the model validation. We discuss the influence of different relative permeability



**Fig. 9** Experimentally measured relative permeability data and curve fits

curves on the characterization, predictions, and average permeability of the core in Sect. 5.3. We refer to these relative permeability curves as follows: curve fit-1 (Kr-1), curve fit-2 (Kr-2), curve fit-3 (Kr-3), and curve fit-4 (Kr-4). For a low injection rate (0.15 cc/min), the experimentally measured remaining brine saturation is approximately 0.55–0.65. In this work, we use the remaining brine saturation of 0.6.

### 3.4 Design of Numerical Simulations

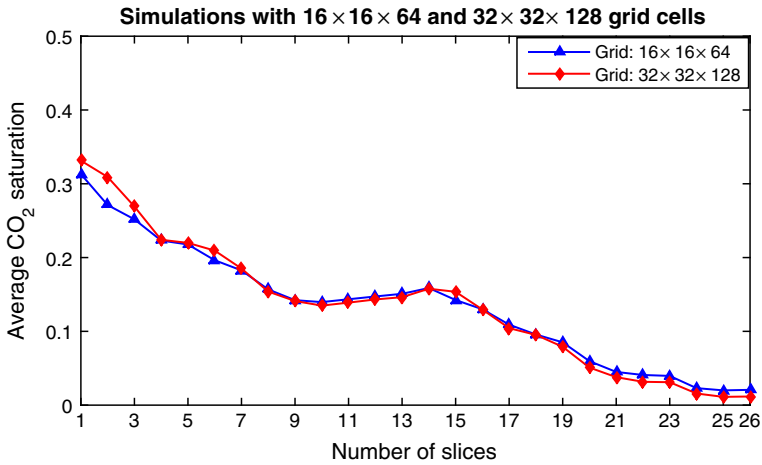
#### 3.4.1 Geological Grid

We use the REV concept to determine a spatial grid on which to assign piecewise constant values of the permeability and porosity fields. We refer to this grid as the geological grid. The REV is the smallest volume at which the volume averages of rock properties do not change significantly. In our work, the porosity-based REV is approximately  $0.78 \times 0.78 \times 0.78 \text{ mm}^3$ , as illustrated in Fig. 8. Typically, the permeability REV is about 2–3 times larger than the porosity REV (Mostaghimi et al. 2013). Therefore, the permeability REV for Berea sandstone is roughly  $2 \times 2 \times 2 \text{ mm}^3$ . In line with this discussion, for discretization of the continuum-scale model, we adopt a spatial mesh size that is comparable to the permeability REV.

#### 3.4.2 Computational Grid

We refer to the spatial grid used to discretize the partial differential equations as the computational grid. We run all compositional simulations on a  $16 \times 16 \times 64$  grid (i.e.,  $2.35 \times 2.35 \times 2.40 \text{ mm}^3$  cell sizes) with some inactive cells to produce a cylindrical core. This mesh size for the computations is consistent with the mesh size of the geological grid.

In principle a computational grid can be finer than REV size, for the purpose of reducing the error between the approximate numerical results and the exact solutions to the partial differential equations. Our numerical simulations show that a finer grid, i.e.,  $32 \times 32 \times 128$ , does not change the saturation profiles significantly. Figure 10 compares slice-averaged  $\text{CO}_2$  saturation for a problem solved on  $16 \times 16 \times 64$  and  $32 \times 32 \times 128$  grids. These are the results for the same problem solved with different mesh refinements. The relative error between the



**Fig. 10** Comparison of the slice-averaged CO<sub>2</sub> saturation values along the length of the core for two computational grids: 16 × 16 × 64 and 32 × 32 × 128

coarse and fine-grid simulations is less than 3%. Since in our characterization/prediction framework we run the forward model thousands of times, we adopted the coarser computational grid in the validation studies reported here.

### 3.4.3 Projection of the Measured Porosity Data on the Computational Grid

The measured porosity has resolution of 250 μm, which is finer than the spatial grid used for simulations. To upscale (or project) the porosity data on the computational grid for the simulations, we use weighted averages of fine-grid porosities to define porosity values on a coarser grid. The weighted average at a cell is defined as:

$$\phi_{\text{cell}} = \frac{1}{v} \sum_{i=1}^{v_n} v_i \phi_i, \tag{8}$$

where  $v_i$  is the volume of voxel  $i$ ,  $\phi_i$  is the porosity value of voxel  $i$ ,  $v = \sum_{i=1}^{v_n} v_i$  is the total volume of voxels, and  $v_n$  is the total number of voxels per cell.

## 4 The Bayesian Markov Chain Monte Carlo Framework

We use a Bayesian framework to quantify uncertainty in the permeability field of a real rock core. The working principle of the Bayesian MCMC framework involves two steps: characterization and predictions, and is described below. This framework has been tested extensively with synthetic computational CO<sub>2</sub>-brine drainage experiments for virtual cores in Akbarabadi et al. (2015). In some respects, these tests are more rigorous than comparisons with physical experiments, since in synthetic tests we have complete knowledge of and control over the uncertain parameters, including the permeability field to be modeled. The framework performs well for all synthetic models of subsurface flows used in the tests. In the case of the synthetic studies reported in Akbarabadi et al. (2015), predictive simulations could be used as a tool to assess the quality of our characterization step. This is not possible

with tests involving experimental data from a real core, because the underlying permeability field is not known.

Within our characterization and prediction steps, we run forward compositional models numerous times, each using a different proposed permeability field. For each permeability proposal, computation of a likelihood function requires a solution of the two-phase compositional model, which, owing to the number of nonlinear flow equations to be solved and the flash calculations, is computationally very intensive. We performed highly resolved simulations using an in-house multiphase compositional simulator (see [Douglas et al. \(2010\)](#) for the object-oriented design of our simulator), in conjunction with MPI and the CUDA parallel computing platforms on a high-performance computer cluster.

## 4.1 The Characterization Step

The characterization step involves two tasks:

1. Find a family of randomly generated permeability fields that serve as inputs to the forward model, such that numerical solutions of this model are in good agreement, within some error, with the measured CO<sub>2</sub> saturation data of scan 1.
2. Determine the suitability of a set of relative permeability curves constructed from the experimental relative permeability data.

Ensemble averages of numerical solutions of the forward model are compared with the measured data of scan 1 for the convenience of visualization. The term ensemble average refers to the average of the simulated CO<sub>2</sub> saturation data for the selected permeability fields from the posterior distribution.

### 4.1.1 The Bayesian Framework: Bayes Rule

In the Bayesian approach, uncertainties in an unknown coefficient in a model are assessed by a posterior distribution. The posterior distribution is proportional to the product of a likelihood function and a prior distribution via Bayes rule:

$$\pi(k) = P(k|d_m) \propto P(d_m|k)P(k). \quad (9)$$

Here  $P(d_m|k)$ ,  $P(k)$ ,  $d_m$ , and  $k$  denote the likelihood function, the prior distribution of the uncertain parameters of the permeability field, the measured data, and the proposed permeability field, respectively. In applying the Bayes rule, due to the high dimensionality and heterogeneity of the permeability field, we use the values of a smaller set of uncertain parameters instead of the permeability values themselves, taking advantage of a decomposition discussed below in Sect. 4.1.3.

We assume that the likelihood function follows a Gaussian distribution ([Efendiev et al. 2005](#); [Ginting et al. 2013](#)):

$$P(d_m|k) \propto \exp\left(-\frac{\|d_m - d_s\|^2}{2\sigma^2}\right), \quad (10)$$

where  $d_s$  is the simulated data when the observed permeability field is  $k$ ,  $\sigma^2$  is the precision parameter associated with the measured and simulated data, and the error norm is given by

$$\|d_m - d_s\|^2 = \sum_{i=1}^{T_s} \left[ \sum_{j=1}^{N_s} \left( \sum_{l=1}^{N_c} (S_m^{i,j,l} - S_s^{i,j,l})^2 \right) \right]. \quad (11)$$

Here,  $T_s$ ,  $N_s$ , and  $N_c$  denote the number of scans used for characterization (in our case  $T_s = 1$ ), the number of slices and the number of cells in a slice, respectively. The variables  $S_m$  and  $S_s$  denote the measured and simulated CO<sub>2</sub> saturation values at transverse slices of the core, respectively.

#### 4.1.2 The Markov Chain Monte Carlo Method (MCMC)

The MCMC method is used to compute numerically the form of a posterior distribution because, in general, such closed form is not available. We use the Metropolis–Hastings MCMC method (Hastings 1970) to sample the permeability field from the posterior distribution. In principle, the algorithm generates a Markov chain with limiting distribution  $\pi(k)$  that replicates the target distribution. In other words, the generated limiting distribution  $\pi(k)$  is invariant. The Metropolis–Hastings MCMC algorithm is given in Algorithm 1.

---

#### Algorithm 1 Metropolis–Hastings MCMC (Hastings 1970)

---

Step 1. Given a permeability proposal  $k_n$  in the Markov chain, generate a permeability proposal  $k$  from the distribution  $q(k|k_n)$  and a random number  $\zeta$  from a uniform distribution  $[0,1]$ .

Step 2. Compute

$$\rho(k_n, k) = \min \left\{ 1, \frac{q(k_n|k)P(d_m|k)P(k)}{q(k|k_n)P(d_m|k_n)P(k_n)} \right\}.$$

Step 3. The next step in the Markov chain is

$$k_{n+1} = \begin{cases} k & \text{if } \rho(k_n, k) \geq \zeta, \\ k_n & \text{otherwise.} \end{cases}$$


---

In Algorithm 1,  $q(k_n|k)$ ,  $k$ , and  $k_n$  represent the proposal distribution, current state of the parameters, and previously accepted state of the parameters, respectively. We use a random walk Metropolis–Hastings algorithm, in which case the proposal distribution is given by

$$\xi_j^{(n+1)}(\omega) = \beta \xi_j^{(n)}(\omega) + \sqrt{1 - \beta^2} \epsilon_n, \quad n = 0, 1, 2, \dots, \tag{12}$$

where  $\beta$  is called the tuning parameter, satisfying  $0 \leq \beta \leq 1$ , and  $\epsilon_n$  is an  $\mathcal{N}(0, 1)$ -random variable (Cotter et al. 2013). The elements  $\xi_j^{(n+1)}(\omega)$  are  $\mathcal{N}(0, 1)$  and are random coefficients in the Karhunen–Loève Expansion (that will be presented in the next section). As stated earlier, due to the heterogeneities and large dimension of the permeability field, its explicit use to compute the prior probability in the MCMC algorithm is not practical. Instead, we use the vector  $(\xi_1(\omega), \dots, \xi_N(\omega))$ . Therefore,  $k_n$  in the Metropolis–Hastings algorithm will be replaced by  $\xi_j^{(n+1)}(\omega)$ —the inferred KLE coefficients. Using a random-walk sampler, the proposal distribution is symmetric and  $q(k_n|k)/q(k|k_n) = 1$  in step 2 of algorithm.

#### 4.1.3 The Karhunen–Loève Expansion

The Karhunen–Loève Expansion (KLE) (Loève 1977; Wong 1971) is a useful technique for approximating high- or infinite-dimensional stochastic processes with a small number of

random variables and has been used by several authors in porous media problems (Efendiev et al. 2005; Douglas et al. 2006; Efendiev et al. 2006; Ma et al. 2008).

In core-scale simulations, the spatial grid for the permeability field can have as many as  $10^5$  cells, and therefore reduction of the space dimensional is computationally helpful. We provide a brief overview of the reduction technique based on the KLE; for more details, we refer to Akbarabadi et al. (2015) and references therein. The truncated KLE is given by

$$Y(\mathbf{x}, \omega) = \sum_{j=1}^N \lambda_j \xi_j(\omega) \psi_j(\mathbf{x}). \quad (13)$$

Here  $Y(\mathbf{x}, \omega)$  is a second-order stochastic process, i.e.,  $Y(\mathbf{x}, \omega) \in L^2(\Omega)$ ,  $\Omega \subseteq \mathbf{R}^3$ ,  $\mathbf{x} = (x, y, z) \in \Omega$ ,  $\omega$  is a random variable,  $N$  denotes total number of dominant terms in the series expansion, and  $\xi_j(\omega)$  belongs to a Gaussian distribution with mean zero and variance one. The deterministic quantities  $\lambda_j$  and  $\psi_j(\mathbf{x})$  are the eigenvalues and the corresponding eigenfunctions of the following equation:

$$\int_{\Omega} R(\mathbf{x}_1, \mathbf{x}_2) \psi_i(\mathbf{x}_2) d\mathbf{x}_2 = \lambda_i \psi_i(\mathbf{x}_1), \quad i = 1, 2, \dots, N, \quad (14)$$

where  $R(\mathbf{x}_1, \mathbf{x}_2)$  is the covariance function given by:

$$R(\mathbf{x}_1, \mathbf{x}_2) = \exp\left(-\frac{|\mathbf{x}_1 - \mathbf{x}_2|}{l}\right). \quad (15)$$

Here  $l$  is the correlation length. The choice of Eq. (15) as a covariance function for the permeability field is motivated by the form of the covariance of the measured porosity field (see Sect. 2.2.3). Figure 7 shows how accurately Eq. (15) with  $l = 0.1$  cm fits the covariance form of the porosity field. As stated earlier, to the best of our knowledge the use of porosity correlation properties to define the prior for absolute permeability modeling has not appeared in the literature, and is used here for the first time.

We assume that the permeability field is distributed log-normally with spatial structure entirely determined by its 2-point statistics. The permeability distribution and the stochastic process are connected by the relation  $k(\mathbf{x}, \omega) = M \exp(sY(\mathbf{x}, \omega))$ , where  $s > 0$  is the heterogeneity strength of the permeability field, and  $M$  is a reference permeability value that we take to be the measured permeability of the core. It can easily be measured in laboratory; in our case it is 612 mD (Akbarabadi and Piri 2013). In contrast, the strength of the permeability field is not known and is taken to be a stochastic parameter to be determined by the available dynamic data, that is,  $\text{CO}_2$  saturation values at transverse slices of scan 1. We remark that we do not make an attempt to recover the value of  $M$  in our study. Instead, our focus is in finding the local permeability field (a function of position) so that our forward-in-time numerical simulations fit the measured data. We refer to the effective core permeability computed from the solution of Laplace's equation as the core average permeability. It will be discussed in Sect. 5.3.

#### 4.1.4 Other Parametrization Approaches

We remark that it would be of interest, and we intend to pursue this line of work in the near future, to compare the parametrization just described with, for instance, the gradual deformation (GD) of Hu et al. (2001) that can reduce the dimensionality of the model without smoothing the permeability maps (as it can happen with a truncated KLE expansion). Combining that technique with FFT based methods (Ravalec et al. 2000) yields fast and



flexible heterogeneous maps generation techniques. This model reduction technique can also be combined with MCMC techniques (Romary 2009). GD avoids the undesirable smoothing effect associated with the KLE approach. However, one may experience low acceptance rates, and recently proposed MCMC acceleration methods (Ginting et al. 2015) may be applicable here.

## 4.2 The Prediction Step

We assess the predictive capability of forward models using Monte Carlo simulations, performed with selected permeability fields from the posterior distribution. For each member of the selected family, we perform a fine-grid simulation for times later than those used in the characterization step. We compute the weighted ensemble average and compare with the measured data of future scans, namely, scans 2–7; see Table 1. The weighted ensemble average is defined as an ensemble average with data for each simulation weighted by the number of realizations rejected between two successive accepted realizations in the MCMC chain. The model is validated if the weighted ensemble average matches the measured data of scans taken after scan 1.

## 5 Validation Results and Discussion

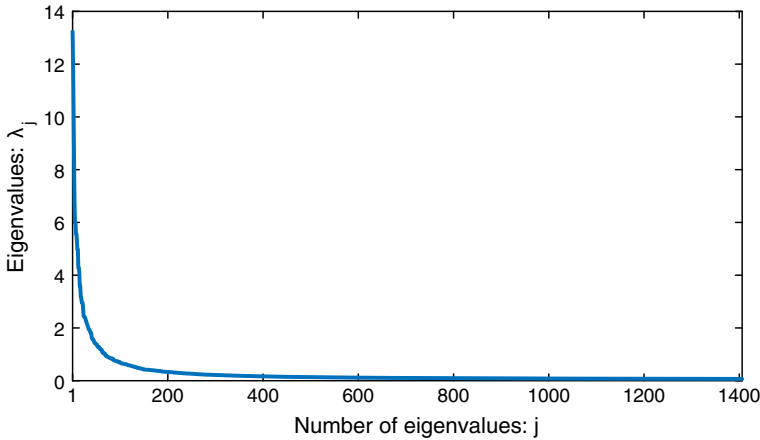
Numerical simulations presented here are performed with four sets of relative permeability curves  $k_{\text{raq}}(S_{\text{aq}})$  and  $k_{\text{rnaq}}(S_{\text{aq}})$ . We use the following notation to refer to simulations with different relative permeability curves:

- Study I: Simulations with relative permeability curve pair (Kr-1,  $k_{\text{rnaq}}$ );
- Study II: Simulations with relative permeability curve pair (Kr-2,  $k_{\text{rnaq}}$ );
- Study III: Simulations with relative permeability curve pair (Kr-3,  $k_{\text{rnaq}}$ );
- Study IV: Simulations with relative permeability curve pair (Kr-4,  $k_{\text{rnaq}}$ ).

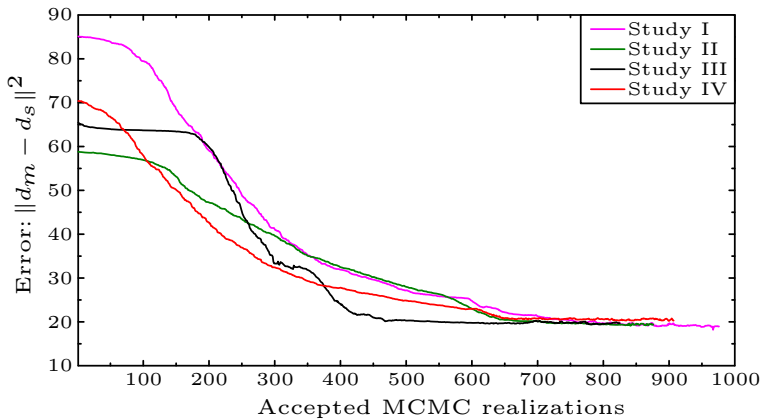
As stated earlier, the numerical simulations are fully consistent with the real experiments. The data used in simulations are given in Sect. 2. The boundary conditions are as follows: the inlet of the core is subject to a constant supercritical CO<sub>2</sub> injection rate of 0.15 cc/min; the outlet pressure is constant at 11 MPa, and no-flow boundary conditions apply at all other boundaries. We consider the first 400 dominant terms in the KLE, i.e.,  $N = 400$ , as illustrated in Fig. 11. In all studies the precision parameter,  $\sigma^2$ , is set to  $4 \times 10^{-3}$ . Dimensional analysis (Hilfer and Øren 1996) shows that the problem considered here is gravity dominated. Thus, for computational efficiency, capillary pressure is not included in the numerical experiments reported below.

### 5.1 The Characterization Results

We now present numerical results related to the characterization step of our Bayesian framework. Figure 12 displays the error between measured and simulated data (Eq. 11) for Studies I, II, III, and IV. Note that the errors decrease until they reach the precision level (determined by  $\sigma^2$  in Eq. 10), indicating the convergence of the chains. The numerical mean value obtained for  $\sigma^2$  is  $3.6 \times 10^{-3}$  which is close to the proposed one ( $4.0 \times 10^{-3}$ ) indicates that the required accuracy was reached. However, this is not the only factor used to evaluate the convergence of a study. Typically, every sampled parameter must reach some equilibrium to call a study convergent. We declare convergence when: (1) the desired precision level has



**Fig. 11** Eigenvalues of the KLE using exponential covariance function

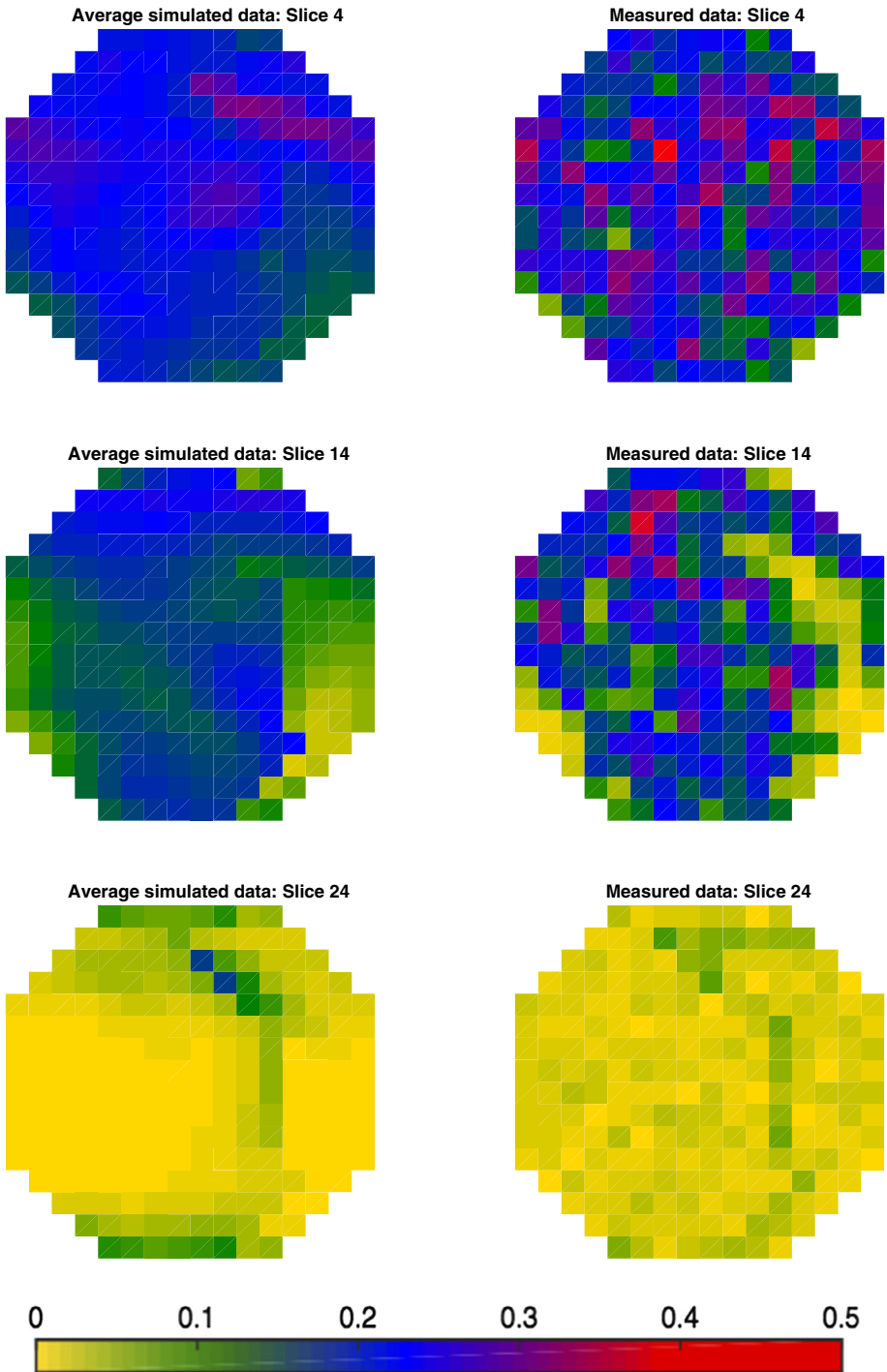


**Fig. 12** Error between the measured and simulated data vs. accepted MCMC realizations

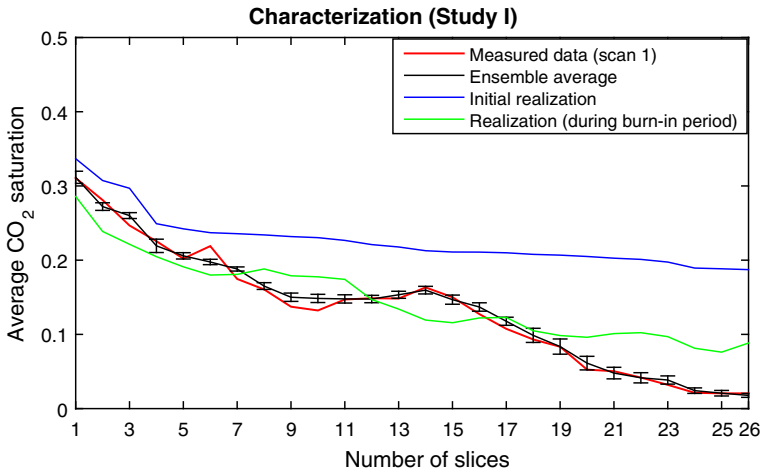
been achieved (see Fig. 12); and (2) the chain for the heterogeneity strength has stabilized (discussed later—see Fig. 18). This study is important in determining the burn-in period for the Markov chains. The burn-in period is defined as an initial period of the selected realizations for which the error between the measured and simulated data has not yet stabilized. After the burn-in period the realizations are sampled from the stationary posterior distribution. On the basis of the errors shown in Fig. 12 and the convergence of the heterogeneity strength illustrated in Fig. 18, we estimate the burn-in periods to be about 480, 550, 650, and 650 initial realizations for Studies I, II, III, and IV, respectively.

Figure 13 shows the measured data (right column) and the ensemble average simulated non-aqueous-phase saturation (left column) at the transverse slices 4, 14, and 24, for study II. Slice 4 is close to the inlet boundary, slice 14 is from the middle, and slice 24 is close to the outlet boundary of the core. The figure shows that the Bayesian framework is able to capture dominant features and general trends of the  $\text{CO}_2$  saturation fields observed in the core.

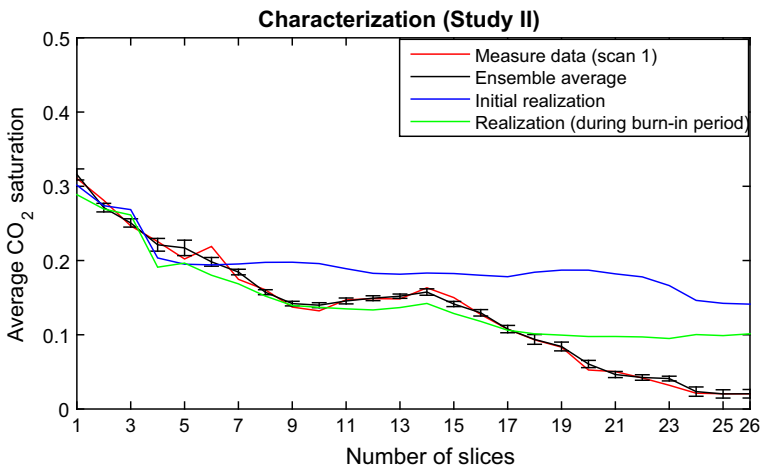
Figures 14 and 15 show the measured slice-averaged  $\text{CO}_2$  saturation and the ensemble average at slices along the length of the core for Studies I and II, respectively. Simulated results



**Fig. 13** (Characterization stage) Comparison between the measured and simulated data at 2D slices for Study II: (Left) Simulated data. (Right) Measured data. First, second, and third rows correspond to slices 4, 14, and 24, respectively



**Fig. 14** (Characterization stage) Slice-averaged CO<sub>2</sub> saturation along the core of the measured data, the initial realization, a realization during the burn-in period and the ensemble average for Study I

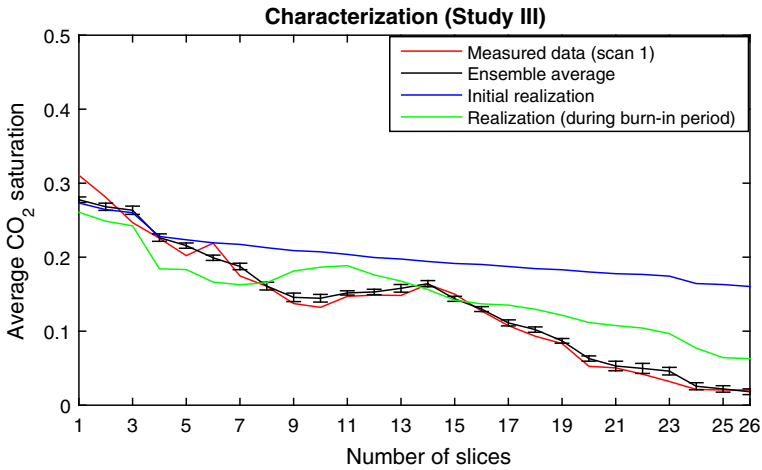


**Fig. 15** (Characterization stage) Slice-averaged CO<sub>2</sub> saturation along the core of the measured data, the initial realization, a realization during the burn-in period and the ensemble average for Study II

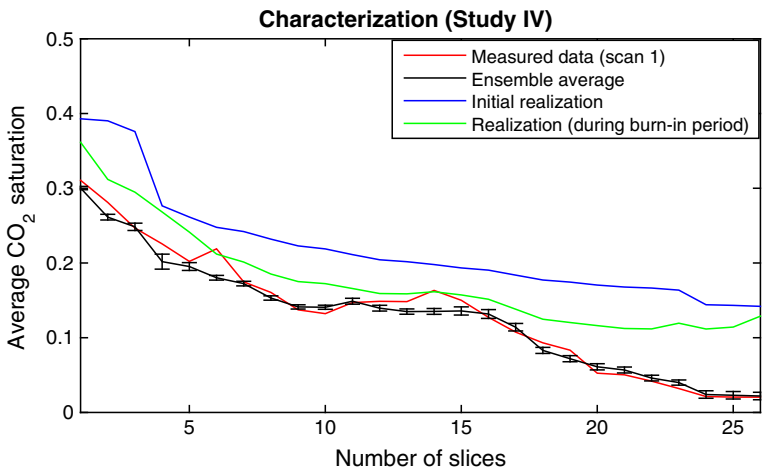
show excellent agreement with the measured data, within some error. These results, along with those displayed in Fig. 19, indicate a successful characterization step. The acceptance rate for permeability proposals is 15–20%.

Figure 16 compares numerical results with the measured data for Study III. Except for a small discrepancy close to the inlet, the simulated and measured curves are in good agreement. We will investigate the predictive capability of this forward model next. Figure 17 plots the ensemble average and the measured data at the slices of the core for Study IV. Again, a good agreement between the simulated and measured data is evident. The error bars represent standard deviation of the data about the mean.

We now turn to a discussion of the determination of the heterogeneity strength using the dynamic (time-dependent CO<sub>2</sub> saturation) data. Figure 18 illustrates convergence of Markov

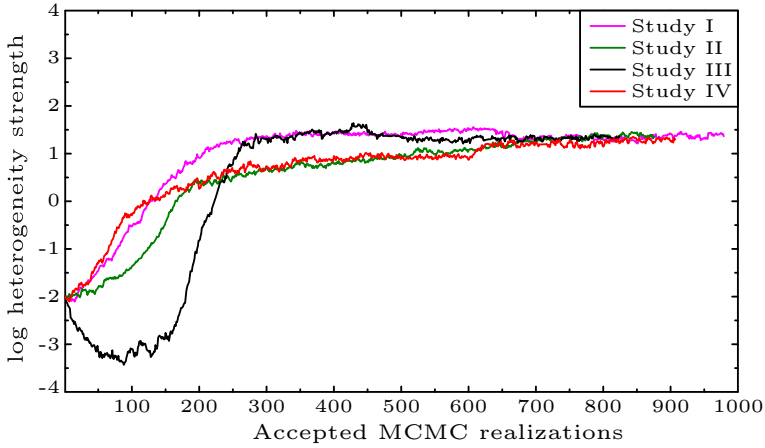


**Fig. 16** (Characterization stage) Slice-averaged CO<sub>2</sub> saturation along the core of the measured data, the initial realization, a realization during the burn-in period and the ensemble average for Study III

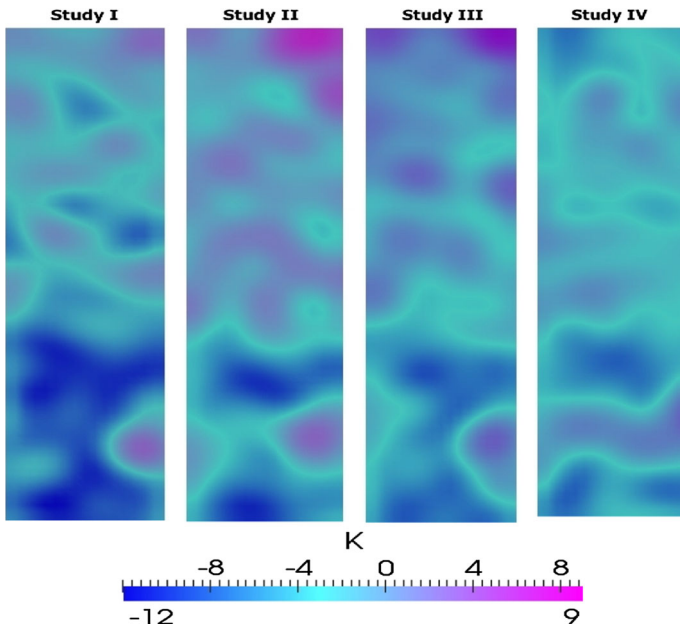


**Fig. 17** (Characterization stage) Slice-averaged CO<sub>2</sub> saturation along the core of the measured data, the initial realization, a realization during the burn-in period and the ensemble average for Study IV

chains for the heterogeneity strength of the permeability field. As discussed earlier, the strength is given by  $s = \exp(\xi^{n+1})$ , where  $\xi^{n+1}$  is  $\mathcal{N}(0, 1)$ -random variable and generated using Eq. (12). Figure 18 shows that the sampled parameter  $\xi^{n+1}$  has reached equilibrium. It is important to point out that the convergence of Markov chains is independent of the initial value as shown in our previous work (Akbarabadi et al. 2015). The oscillations in strength values are very small; they stabilize in a neighborhood of the value 1.3 (on a logarithmic scale) for all studies. However, they do not converge exactly to the same number. This fact is related to the distinct average permeability of the realizations from the posterior distribution that is discussed below. The framework used here can reveal the strength of the permeability field. Our synthetic computational experiments (Akbarabadi et al. 2015) show that the heterogeneity strength converges to the reference strength and is independent of the initial



**Fig. 18** Illustration of the convergence of the Markov chain for the heterogeneity strength of the permeability field vs. accepted MCMC realizations for Studies I, II, III, and IV. The vertical axis displays the random number,  $\xi^{n+1}$ , generated to compute the strength given by  $s = \exp(\xi^{n+1})$ . The convergence is independent of the initial value as shown in our previous work (Akbarabadi et al. 2015)



**Fig. 19** (Left-Right) An illustration of vertical cuts of permeability fields (on a logarithmic scale) of the core for Studies I, II, III, and IV. These fields belong to the posterior distribution

strength. In Fig. 19, we show some two-dimensional vertical cuts of selected permeability fields for Studies I, II, III, and IV. Although there is no reason to expect similarities among these fields (it is an ill-posed problem), some common trends were observed, probably due to the large CO<sub>2</sub> saturation data used in the likelihood function. We have 26 transverse slices and 216 measured CO<sub>2</sub> saturation values per slice, so the total number of measurements used

in the likelihood is 5616. We close this subsection with the discussion of signal-to-noise ratio (SNR). The SNR is defined as the ratio of the root mean squared error (RMSE) of the initial realization to the average posterior RMSE. Mathematically,

$$\text{SNR} = \frac{\|d_m - d_s\|_{\text{init}}}{\sqrt{N}} \left( \frac{\langle \|d_m - d_s\|_{\text{post}}^i \rangle}{\sqrt{N}} \right)^{-1} = \frac{\|d_m - d_s\|_{\text{init}}}{\langle \|d_m - d_s\|_{\text{post}}^i \rangle} \quad (16)$$

Here  $N$  is the total number of measurements employed in the likelihood, the angle brackets represent the average, the superscript  $i$  is from 1 to the size of the posterior sample space, and the subscripts *init* and *post* denote the initial and posterior realizations, respectively. In our studies, the SNR is approximately 2.0 which indicates extraction of useful information from the posterior (Mao et al. 2007).

## 5.2 Prediction Results

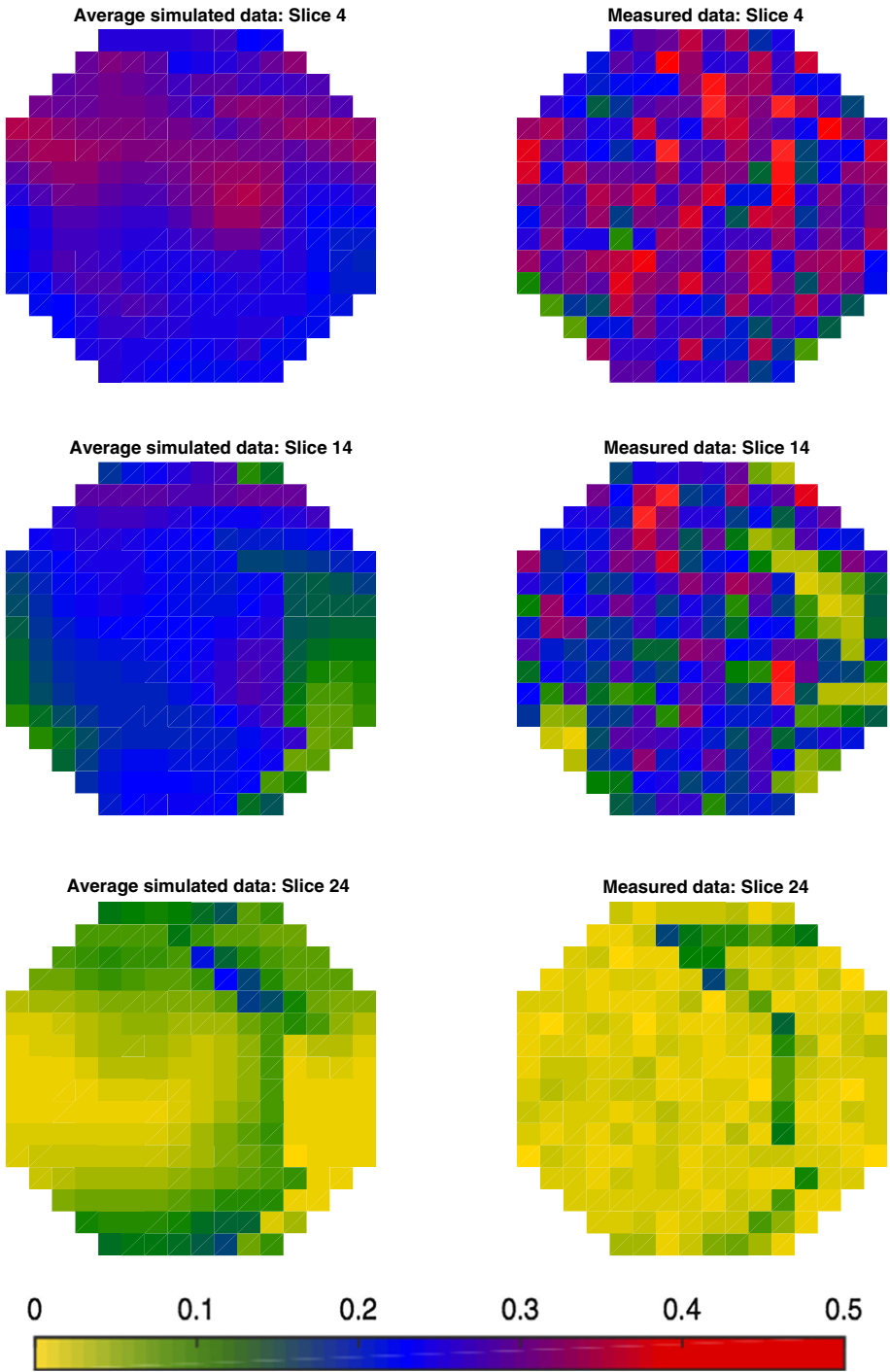
The aim of predictive simulations is to validate forward models, which consist of the compositional two-phase flow model along with specific relative permeability curves. Figure 20 shows the measured data (right column) and the weighted ensemble average simulated CO<sub>2</sub> saturation (left column) at the transverse slices 4, 14, and 24, for study II. Similar to the characterization results, the simulation results at the prediction stage possess the dominant features and general trends of the observed CO<sub>2</sub> saturation in the core. These results indicate that capturing general trends of the permeability field are enough to produce accurate predictions.

Figures 21, 22 and 23 show a comparison between the weighted ensemble average and the measured data of scans 3, 5, and 7, respectively, for Studies I, II, III, and IV. We observe excellent quantitative agreement of the simulated data with the available experimental data for Studies I and II. Thus, these forward models can be used for accurate predictions.

Predictive simulations for Study III inherit the discrepancies observed in the characterization stage, and as a result validation is not established. It can be seen in Fig. 23 that numerical results for Study IV do not match the measured data very well. That is, for Study IV the model is not validated successfully even though we see a good agreement of the simulated data with the measured data at the characterization step.

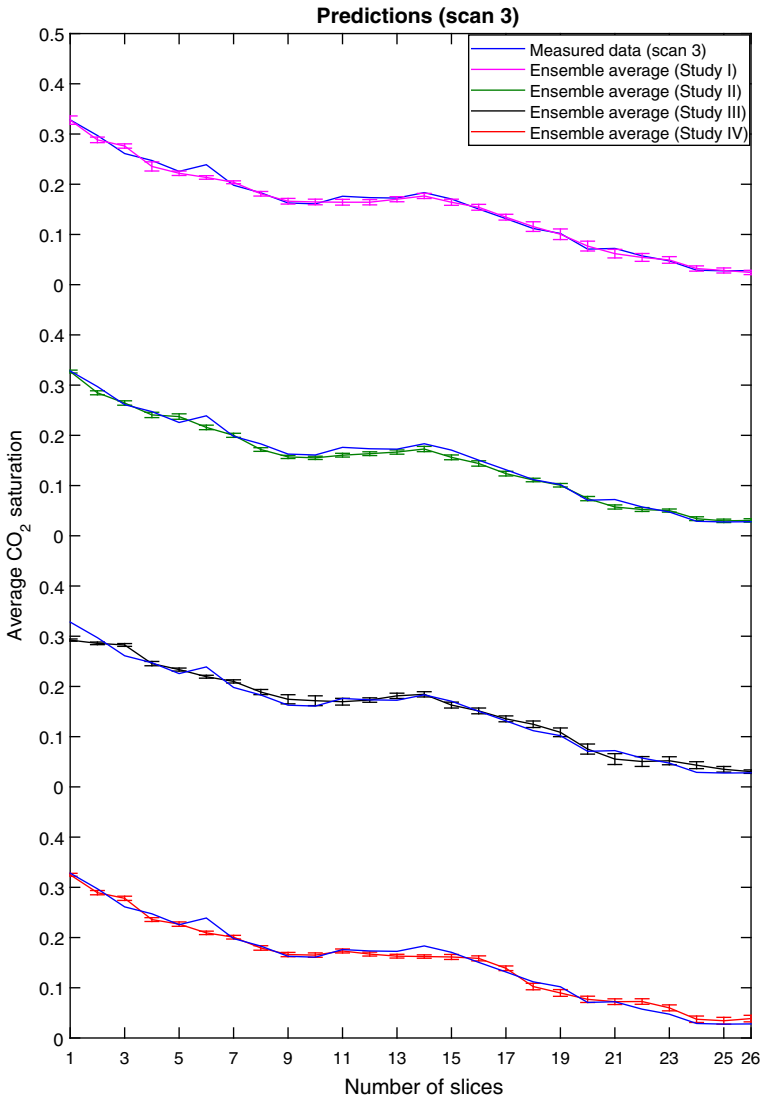
Table 4 displays the root mean square differences (RMSD) between the measured and simulated data corresponding to Figs. 21, 22 and 23. Note that the error is growing with time in all studies; however, this growth is considerably larger in studies III and IV. This conclusion is further illustrated in Fig. 24 which depicts the percentage increase in the RMSD in the prediction stage. This is the increase with respect to the RMSD of the ensemble average in the characterization stage. Finally, we plot the RMSD of scan 7 in Fig. 25 that also shows larger differences for studies III and IV.

As alluded to earlier, the mapping between the measured relative permeability data (discrete points) and relative permeability functions (differentiable curves) is not unique. This fact brings additional uncertainty to mathematical models of multiphase flow. Our study above indicates that adequate forward models can be determined within a Bayesian framework. However, it would be desirable to use the available time-dependent saturation data to determine the relative permeability functions, which can be written in terms of a small number of additional stochastic parameters. A synthetic study to quantifying uncertainty in flow functions is reported in Subbey et al. (2006). The extension of the Bayesian framework of Akbarabadi et al. (2015) to select acceptable permeability samples as well as relative permeability curves is a promising avenue for further research.



**Fig. 20** (Prediction stage scan 7) Comparison between the measured and simulated data at 2D slices for Study II: (Left) Simulated data. (Right) Measured data. First, second, and third rows correspond to slices 4, 14, and 24, respectively

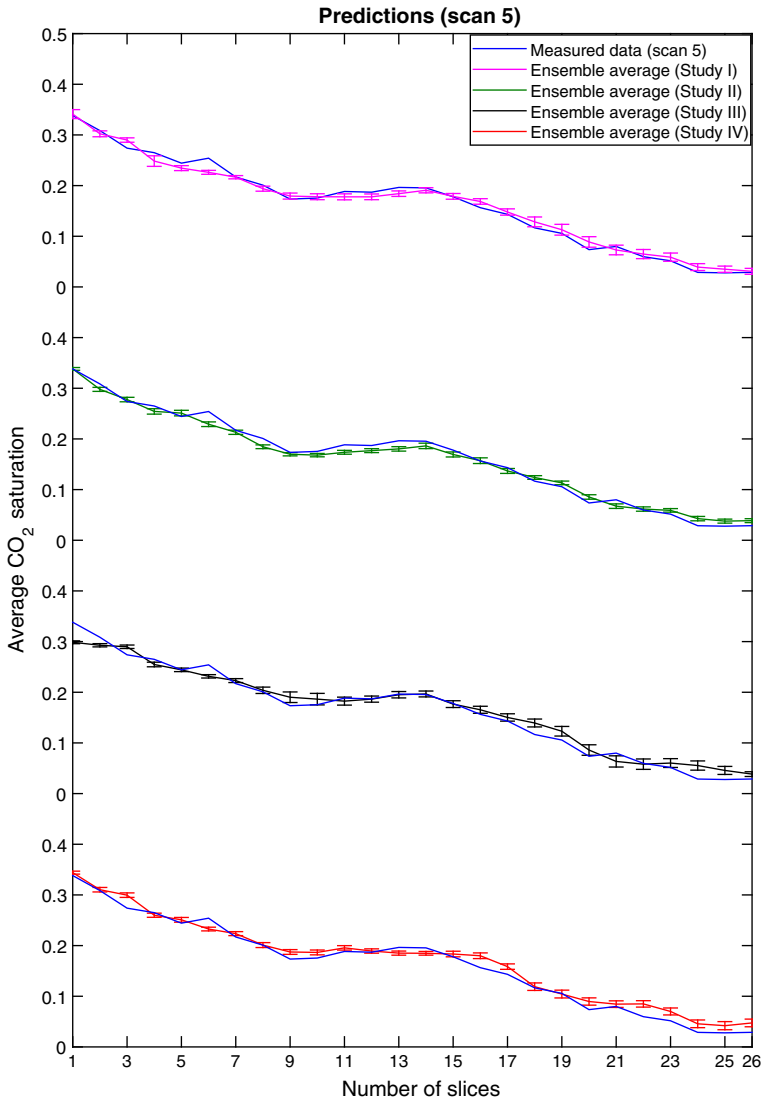




**Fig. 21** (Prediction stage: scan 3) Comparison between the slice-averaged CO<sub>2</sub> saturation along the core of the measured data and the ensemble average of simulated data for all studies

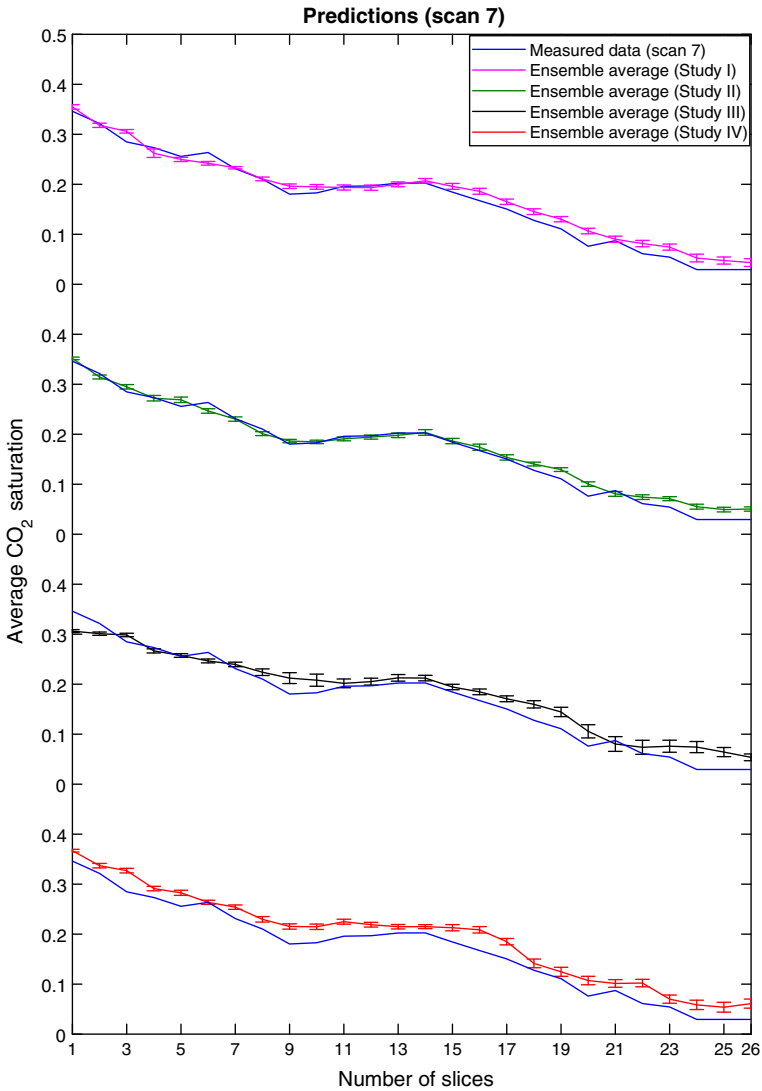
### 5.3 Discussion of Average Permeability

We now investigate the effect of using different relative permeability curves on the predicted average permeability of the core. For each selected permeability field, the average permeability of the core is computed by solving the pressure equation for single-phase flow. Figure 26 illustrates the average permeability of the accepted MCMC realizations for Studies I, II, III, and IV. The arithmetic average of the average permeability fields for the selected families is about 55, 65, 180, and 250 mD for Studies I, II, III, and IV, respectively. Both the inferred heterogeneity strength and the fixed relative permeability curve affect the inferred average



**Fig. 22** (Prediction stage: scan 5) Comparison between the slice-averaged CO<sub>2</sub> saturation along the core of the measured data and the ensemble average of simulated data for all studies

permeability. In line with the Bayesian framework of Akbarabadi et al. (2015) the reference permeability  $M$  has been fixed. However, it may be of interest to infer this value from the data and we intend to investigate this possibility. We remark that “a pressure” is also measured in two positions of the core (see the pressure ports in Fig. 1). However, this pressure is not a phase pressure (they cannot be directly measured with existing experimental tools) and we also intend to investigate how to take advantage of these measurements to reduce uncertainty.



**Fig. 23** (Prediction stage: scan 7) Comparison between the slice-averaged CO<sub>2</sub> saturation along the core of the measured data and the ensemble average of simulated data for all studies

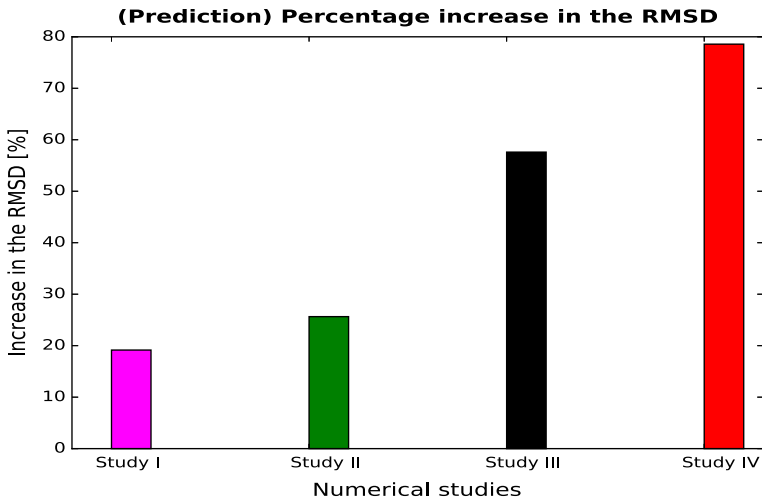
### 5.4 Additional Remarks

When the validation is established, simulated data have the same general trends as that of the experimental data, even though we see slightly more CO<sub>2</sub> from simulations toward the bottom of the core. Uncertainty in the measured data (experimental measurement errors) might be a possible reason for a slight discrepancy close to the outlet.

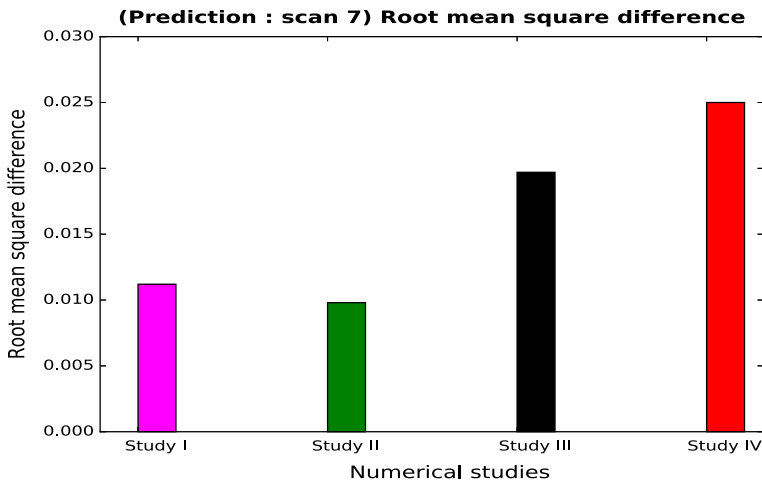
To construct relative permeability curves, we use best curve fits to the experimentally measured relative permeability data. Our initial numerical studies (not reported here) were based on a best curve fit that gives minimum error with the experimental relative perme-

**Table 4** (Prediction stage) RMSD between the measured and simulated data

| Study | Scan 3 | Scan 5 | Scan 7 |
|-------|--------|--------|--------|
| I     | 0.0075 | 0.0089 | 0.0112 |
| II    | 0.0070 | 0.0091 | 0.0098 |
| III   | 0.0122 | 0.0136 | 0.0197 |
| IV    | 0.0109 | 0.0132 | 0.025  |

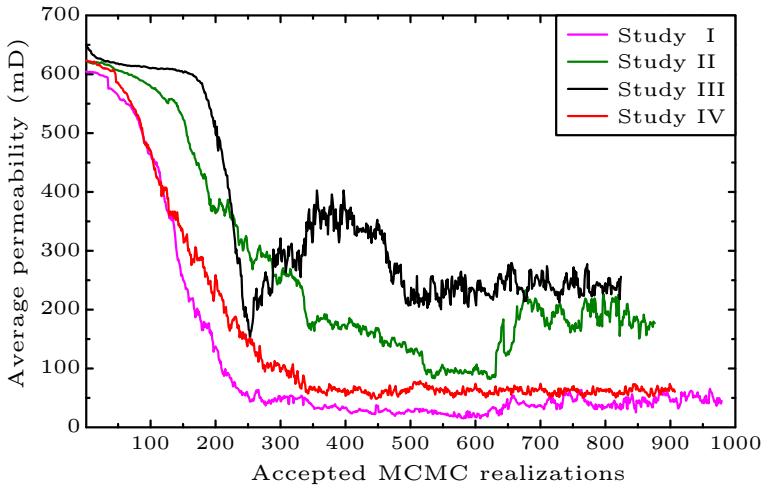


**Fig. 24** (Prediction stage) An illustration of the percentage increase in the root mean square difference of the measured and simulated data in reference to the ensemble average at the characterization stage



**Fig. 25** (Prediction stage : scan 7) The root mean square error difference of the measured and simulated data

ability data and the remaining brine saturation was 0.47. This curve was used in the study reported in [Rahunanthan et al. \(2014\)](#). However, we could not validate models for such relative permeability curves because the prediction step failed. Upon further discussion with the



**Fig. 26** Shown is the average permeability of the selected realizations for Studies I, II, III, and IV

experimental scientists in our group, it was indicated that there is uncertainty in the determination of the remaining brine saturation, and a higher value for this parameter should be tested. Thus, in the studies reported here we have considered a remaining brine saturation of 0.6 along with distinct tolerances for the fitting procedure. We believe it would be of interest to perform a sensitivity study on the remaining brine saturation to better understand the influence of this parameter. This study is, however, outside the scope of this work. Due to the slow convergence of the MCMC scheme we also intend to apply adaptive algorithms that automatically tuning the proposed distribution aiming at significantly improving the efficiency of the MCMC method (Haario et al. 1999, 2001, 2006; Vrugt et al. 2009; Vrugt and Braak 2011). Note that such algorithms can be associated with other techniques of dimensionality reduction (such as using the interpolated Gaussian process generation via circulant embedding (Dietrich and Newsam 1997; Laloy et al. 2015)).

## 6 Conclusions

We assessed the validity of a compositional two-phase flow model for CO<sub>2</sub>-brine flow at the laboratory scale. The primary source of uncertainty one has to overcome to be able to perform predictive numerical simulations is in the determination of the 3D permeability field of the core.

In this study, we employed a Bayesian Markov chain Monte Carlo framework of Akbarabadi et al. (2015) to validate a compositional model describing an unsteady-state CO<sub>2</sub>-brine drainage experiment, at the laboratory scale, performed at aquifer conditions. We claimed model validation when: (1) numerical results showed that the Bayesian framework has accurately characterized the core, and (2) CO<sub>2</sub> saturations in Monte Carlo predictive simulations were in good agreement with measured data. Characterization refers to selecting a family of permeability fields such that solutions of the forward model (consisting of a discretized system of partial differential equations along with relative permeability curves) match the measured dynamic data (CO<sub>2</sub> saturation at slices).

Large-scale numerical simulation studies show that forward models can be validated by this Bayesian framework. Moreover, the heterogeneity strength of the permeability field is considered as a stochastic parameter to be determined by the dynamic data, and our results reveal the convergence of this parameter. Also, to assess the effects of nonuniqueness in the mapping between relative permeability data (discrete data points) to relative permeability curves, we examined the results generated using different relative permeability curves. Our numerical results show that model validation is subject to the choice of relative permeability curves. Moreover, we found that this choice may affect the average permeability of the selected samples of the permeability field of the core.

The work presented here has identified some open problems that we intend to address. One important development would be the extension of the Bayesian framework of Akbarabadi et al. (2015) to identify appropriate relative permeability curves; see Subbey et al. (2006). In another direction, one could also consider the generation of permeability samples with average permeability close to the experimental value.

**Acknowledgements** The authors are grateful to Prof. Myron B. Allen for his suggestions and for carefully reading the manuscript. F. P. and M. P. were partially supported by DOE grant DE-FE0004832 and the Clean Coal Technologies Research Program of the School of Energy Resources of the University of Wyoming (1100 20352 2012). F. P. was also funded in part by NSF-DMS 1514808, a Science Without Borders/CNPq-Brazil grant and UT Dallas. M. P. was also funded in part by Hess Corporation and the School of Energy Resources at the University of Wyoming. M. B. was funded by CNPq-Brazil.

## 7 Appendix A: Relative Permeability Analytic Parameterization

The analytical parameterizations of the relative permeability curves presented in Fig. 9 are listed below:

$$\text{Kr-1}(S_{\text{aq}}) = \exp(-36.1760530 + 35.906092\sqrt{S_{\text{aq}}}) \quad (17)$$

$$\text{Kr-2}(S_{\text{aq}}) = \exp(-38.46 + 36.0\sqrt{S_{\text{aq}}} + 1.6S_{\text{aq}}^2 + 0.2S_{\text{aq}}) \quad (18)$$

$$\text{Kr-3}(S_{\text{aq}}) = \exp(-37.1760530 + 36.3106092\sqrt{S_{\text{aq}}}) \quad (19)$$

$$\text{Kr-4}(S_{\text{aq}}) = \frac{-0.0108 + 0.0198S_{\text{aq}}}{1.0 - 0.9832S_{\text{aq}}} \quad (20)$$

For reference, these aqueous phase relative permeability functions correspond to the relative permeability curves used in Study I, Study II, Study III, and Study IV, respectively, and the non-aqueous relative permeability function is given below:

$$k_{\text{rnaq}}(S_{\text{aq}}) = -1.7212298S_{\text{aq}}^{10} + 3.9501271S_{\text{aq}}^8 - 3.5647535S_{\text{aq}}^6 + 2.2058758S_{\text{aq}}^4 \\ - 1.3020969S_{\text{aq}}^2 + 0.4320773$$

## References

- Akbarabadi, M., Borges, M., Jan, A., Pereira, F., Piri, M.: A bayesian framework for the validation of models for subsurface flows: synthetic experiments. *Comput. Geosci.* **19**(6), 1231–1250 (2015)
- Akbarabadi, M., Piri, M.: Relative permeability hysteresis and capillary trapping characteristics of supercritical CO<sub>2</sub>/brine system: an experimental study at reservoir conditions. *Adv. Water Resour.* **52**, 190–206 (2013)
- Bear, J.: *Hydraulics of Groundwater*. McGraw-Hill, New York (1979)

- Benson, S.M., Tomutsa, L., Silin, D., Kneafsey, T., Miljkovic, L.: Core scale and pore scale studies of carbon dioxide migration in saline formations. In: Proceedings of the 8th International Conference on Greenhouse GasControl Technologies, IEA Greenhouse Gas Program, Trondheim, Norway (2006)
- Chen, Z., Zhang, Y.: Development, analysis and numerical tests of a compositional reservoir simulator. *Int. J. Numer. Anal. Model.* **5**, 86–100 (2008)
- Cotter, S., Roberts, G., Stuart, A., White, D.: MCMC methods for functions: modifying old algorithms to make them faster. *Stat. Sci.* **28**, 424–446 (2013)
- Dietrich, C.R., Newsam, G.N.: Fast and exact simulation of stationary gaussian processes through circulant embedding of the covariance matrix. *SIAM J. Sci. Comput.* **18**(4), 1088–1107 (1997)
- Douglas, C., Efendiev, Y., Ewing, R., Ginting, V., Lazarov, R.: Dynamic data driven simulations in stochastic environments. *Computing* **77**, 321–333 (2006)
- Douglas, C., Furtado, F., Ginting, V., Mendes, M., Pereira, F., Piri, M.: On the development of a high-performance tool for the simulation of CO<sub>2</sub> injection into deep saline aquifers. *Rocky Mt. Geol.* **45**, 151–161 (2010)
- Efendiev, Y., Datta-Gupta, A., Ginting, V., Ma, X., Mallick, B.: An efficient two-stage Markov chain Monte Carlo method for dynamic data integration. *Water Resour. Res.* **41**, W12423 (2005)
- Efendiev, Y., Hou, T., Luo, W.: Preconditioning Markov chain Monte Carlo simulations using coarse-scale models. *SIAM J. Sci. Comput.* **28**, 776–803 (2006)
- Ginting, V., Pereira, F., Rahunanthan, A.: Rapid quantification of uncertainty in permeability and porosity of oil reservoirs for enabling predictive simulation. *Math. Comput. Simul.* **99**, 139–152 (2013)
- Ginting, V., Pereira, F., Rahunanthan, A.: Multi-physics Markov chain Monte Carlo methods for subsurface flows. *Math. Comput. Simul.* **118**, 224–238 (2015)
- Glimm, J., Sharp, D.H.: Prediction and the quantification of uncertainty. *Phys. D* **133**(1–4), 152–170 (1999)
- Haario, H., Saksman, E., Tamminen, J.: Adaptive proposal distribution for random walk metropolis algorithm. *Comput. Stat.* **14**(3), 375–395 (1999)
- Haario, H., Laine, M., Mira, A., Saksman, E.: DRAM: efficient adaptive MCMC. *Stat. Comput.* **16**(4), 339–354 (2006)
- Haario, H., Saksman, E., Tamminen, J.: An adaptive metropolis algorithm. *Bernoulli* **7**(2), 223–242 (2001)
- Hastings, W.K.: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**, 97–109 (1970)
- Hilfer, R., Øren, P.E.: Dimensional analysis of pore scale and field scale immiscible displacement. *Transp. Porous Media* **22**, 53–72 (1996)
- Hu, L.Y., Blanc, G., Noetinger, B.: Gradual deformation and iterative calibration of sequential stochastic simulations. *Math. Geol.* **33**(4), 475–489 (2001)
- Kong, X., Delshad, M., Wheeler, M.F.: High resolution simulations with a compositional parallel simulator for history matching laboratory CO<sub>2</sub>/brine core flood experiment. *Soc. Pet. Eng. J.* **1**, 664–676 (2014)
- Krause, M., Krevor, S., Benson, S.M.: A procedure for the accurate determination of sub-core scale permeability distributions with error quantification. *Transp. Porous Med.* **98**, 565–588 (2013). doi:[10.1007/s11242-013-0161-y](https://doi.org/10.1007/s11242-013-0161-y)
- Krause, M., Perrin, J.C., Benson, S.M.: Modeling permeability distributions in a sandstone core for history matching coreflood experiments. *Soc. Pet. Eng. J.* **126340**(16), 768–777 (2011)
- Laloy, E., Linde, N., Jacques, D., Vrugt, J.: Probabilistic inference of multi-gaussian fields from indirect hydrological data using circulant embedding and dimensionality reduction. *Water Resour. Res.* **6**(51), 4224–4243 (2015)
- Le Ravalec, M., Noetinger, B., Hu, L.Y.: The FFT moving average (FFT-MA) generator: an efficient numerical method for generating and conditioning gaussian simulations. *Math. Geol.* **32**(6), 701–723 (2000)
- Leal, A.M.M.: Flash equilibrium method for CO<sub>2</sub> and H<sub>2</sub>S storage in brine aquifers with parallel GPU implementation. Master's thesis, Math Department University of Wyoming (2010)
- Loève, M.: Probability Theory. Springer, Berlin (1977)
- Ma, X., Al-Harbi, M., Datta-Gupta, A., Efendiev, Y.: An efficient two-stage sampling method for uncertainty quantification in history matching geological models. *Soc. Pet. Eng. J.* (2008). doi:[10.2118/102476-PA](https://doi.org/10.2118/102476-PA)
- Mao, X., Amini, P., Farhang-Boroujeny, B.: Markov chain Monte Carlo mimo detection methods for high signal-to-noise ratio regimes. In: Global Telecommunications Conference, GLOBECOM'07. IEEE, pp. 3979–3983 (2007)
- Mostaghimi, P., Blunt, M.J., Bijeljic, B.: Computations of absolute permeability on micro-CT images. *Math. Geosci.* **45**, 103–125 (2013)
- Nelson, P.: Permeability–porosity relationships in sedimentary rocks. *Log Anal.* **35**, 38–62 (1994)
- Ovaysi, S., Piri, M.: Direct pore-level modeling of incompressible fluid flow in porous media. *J. Comput. Phys.* **229**, 7456–7476 (2010)

- Perrin, J.C., Benson, S.: An experimental study on the influence of sub-core scale heterogeneities on CO<sub>2</sub> distribution in reservoir rocks. *Transp. Porous Media* **82**, 93–109 (2010)
- Qin, G.: Numerical solution techniques for compositional model. Ph.D. Thesis, Department of Chemical and Petroleum Engineering, University of Wyoming (1995)
- Rahunanthan, A., Furtado, F., Marchesin, D., Piri, M.: Hysteretic enhancement of carbon dioxide trapping in deep aquifers. *Comput. Geosci.* **18**, 899–912 (2014)
- Romary, T.: Integrating production data under uncertainty by parallel interacting Markov chains on a reduced dimensional space. *Comput. Geosci.* **13**(1), 103–122 (2009)
- Shi, J.Q., Xue, Z., Durucan, D.: History matching of CO<sub>2</sub> core flooding CT scan saturation profiles with porosity dependent capillary pressure. *Energy Proc.* **1**, 3205–3211 (2009)
- Shi, J.Q., Xue, Z., Durucan, S.: Supercritical CO<sub>2</sub> core flooding and imbibition in Tako sandstone-influence of sub-core scale heterogeneity. *Int. J. Greenhouse Gas Control* **5**, 75–87 (2011)
- Subbey, S., Monfared, H., Christie, M., Sambridge, M.: Quantifying uncertainty in flow functions derived from SCAL data. *Transp. Porous Media* **65**, 265–286 (2006)
- Trangenstein, J.A., Bell, J.B.: Mathematical structure of compositional reservoir simulation. *SIAM J. Sci. Stat. Comput.* **10**, 817–845 (1989)
- Vrugt, J.A., Ter Braak, C.J.F.: Dream(*D*): an adaptive Markov chain Monte Carlo simulation algorithm to solve discrete, noncontinuous, and combinatorial posterior parameter estimation problems. *Hydrol. Earth Syst. Sci.* **15**(12), 3701–3713 (2011)
- Vrugt, J.A., ter Braak, C.J.F., Diks, C.G.H., Robinson, B.A., Hyman, J.M., Higdon, D.: Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int. J. Nonlinear Sci. Numer. Simul.* **10**(3), 273–290 (2009)
- Wong, E.: *Stochastic Processes in Information and Dynamical Systems*. McGraw-Hill, New York (1971)