HISTORY MATCHING OF SURFACTANT-POLYMER FLOODING

A Thesis

Submitted to the Faculty

of

Purdue University

by

Pratik K. Naik

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Mechanical Engineering

December 2018

Purdue University

West Lafayette, Indiana

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF THESIS APPROVAL

Dr. Arezoo M. Ardekani, Chair

     School of Mechanical Engineering

Dr. Ivan C. Christov

     School of Mechanical Engineering

Dr. Sadegh Dabiri

     School of Mechanical Engineering and Department of Agricultural and Biological Engineering

Approved by:

     Dr. Jay Gore

          Head of the School Graduate Program

To my family.

ACKNOWLEDGMENTS

I had the most amazing two and half years at Purdue and it was all possible because of the immense support and encouragement I received from everyone around here. This thesis is the blend of all these incredible experiences I had at Purdue where I had this incredible opportunity to learn from some of the most influential minds and I would like to express my deepest gratitude to each one of them.

First of all, my advisor Dr. Arezoo Ardekani for her continuous support and encouragement since the day I began this journey. Without her guidance, constant motivation and emotional as well as financial support, this thesis wouldn't have been possible.

To my fellow lab members, in particular Soroush Aramideh for all the brainstorming sessions and interesting discussions we had over these years. Soroush has greatly helped shaped the structure of this thesis work.

I am also thankful to Dr. Rituaj Borgohain and Dr. Nathan Schultheiss from the Enhanced Oil Recovery Lab at Bindley Bioscience Center, Purdue University for generously providing the experimental data to support our numerical findings.

I would also like to acknowledge Dr. Ivan Christov and Dr. Sadegh Dabiri, for accepting to be on my defense committee. I appreciate their invaluable time, interest and insightful feedbacks.

I shall forever be grateful to School of Mechanical Engineering, Purdue University for providing Graduate Assistantship to support my stay here.

Last but not the least, I would like to thank my parents: my mother Kanchan and my father Kiran, for making all the sacrifices, providing unconditional support and giving me the freedom to pursue my dreams.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Figure   Page

SYMBOLS

M	mobility ratio
Ca	capillary number
σ_{ow}	interfacial tension between oil and water
μ_w, μ_n	viscosity of wetting (aqueous) and non-wetting (oleic) phase
λ_w, λ_n	mobility of wetting and non-wetting phase
k_{rw}, k_{rn}	relative permeability of wetting and non-wetting phase
S_w, S_n	saturation of wetting and non-wetting phase
u_l	phase field velocity (water $l = w$, oil $l = n$)
ϕ	porosity
ρ_l	density (water $l = w$, oil $l = n$)
p_c	capillary pressure
p_l	phase pressure (water $l = w$, oil $l = n$)
n_l	Coreys exponent
α_k	dispersivity (longitudinal $k = L$, transverse $k = T$)
μ_p	polymer viscosity
$\dot{\gamma}$	shear rate
R_F	residual resistance factor
$P(\cdot)$	PCE-proxy
y_α	coefficients of PCE-proxy
ξ	vector of model parameters
N_s	size of experimental design set
Ψ_α	multivariate orthonormal polynomial basis
$\mathbb{E}[\cdot]$	expectation of $[\cdot]$
$\mathbb{V}[\cdot]$	variance of $[\cdot]$

ϵ_{mean}^{LOO}	leave-one-out error
ϵ_{norm}^{LOO}	normalized empirical error
L_2	relative Euclidean norm
N_{gen}	Genetic algorithm generation
NPV	net present value
$p(\cdot)$	probability density function of $(\cdot)$
$\hat{\mathbf{x}}$	optimized design variables

ABBREVIATIONS

BCF   berea coreflood

BGO   bayesian global optimization

BHM   bayesian history matching

BL   Buckley-Leverett

CDC   capillary desaturation curve

cEOR   chemical enhanced oil recovery

CMC   critical micelle concentration

EEIHV extended expected improvement over dominated hypervolume

EOR   enhanced oil recovery

EWF   extended water flooding

GA   Genetic algorithm

GP   Gaussian process

HTDS   high total dissolved solids

IAF   information acquisition function

IFT   interfacial tension

IWF   initial water flooding

LHS   latin hypercube sampling

LOO   leave one out

MCMC Markov chain Monte Carlo

MOO   multi-objective optimization

OOIP   original oil in place

PCE   polynomial chaos expansion

PCE-K polynomial chaos expansion- kriging

PDF   probability density function

QoIs	quantities of interest
RRF	residual resistance factor
SA	sensitivity analysis
SP	surfactant-polymer
UQ	uncertainty quantification

ABSTRACT

Naik, Pratik K. M.S.M.E., Purdue University, December 2018. History Matching of Surfactant-polymer Flooding. Major Professor: Arezoo Ardekani.

This thesis presents a framework for history matching and model calibration of surfactant-polymer (SP) flooding. At first, a high-fidelity mechanistic SP flood model is constructed by performing extensive lab-scale experiments on Berea cores. Then, incorporating Sobol based sensitivity analysis, polynomial chaos expansion based surrogate modelling (PCE-proxy) and Genetic algorithm based inverse optimization, an optimized model parameter set is determined by minimizing the miss-fit between PCE-proxy response and experimental observations for quantities of interests such as cumulative oil recovery and pressure profile. The epistemic uncertainty in PCE-proxy is quantified using a Gaussian regression process called Kriging. The framework is then extended to Bayesian calibration where the posterior of model parameters is inferred by directly sampling from it using Markov chain Monte Carlo (MCMC). Finally, a stochastic multi-objective optimization problem is posed under uncertainties in model parameters and oil price which is solved using a variant of Bayesian global optimization routine.

1. SURFACTANT-POLYMER FLOODING FOR ENHANCED OIL RECOVERY

1.1 Enhanced oil recovery using chemical flooding

Enhanced oil recovery (EOR) aims at using different techniques such as chemical flooding, CO_2 injection or thermal recovery to enhance mobilization of trapped oil in reservoirs [1]. The methods available for recovering oil from reservoirs can be broadly classified into three categories: 1) Primary recovery 2) Secondary recover and 3) Tertiary recovery. During the primary recovery process, production wells are drilled into the reservoir and using the natural pressure of the reservoir or gravity in combination with the artificial lift generated by the pumps, the crude oil is driven out to the surface. However, only about 10% of original oil in place (OOIP) could be extracted using this method. During secondary recovery, injection wells are drilled at suitable locations and water is injected through them to build up an artificial differential pressure sweeping the trapped oil. Secondary recovery leads to a production of 20-40% of OOIP. At such a low residual oil saturation, oil is immobile as the differential pressure alone cannot overcome the capillary pressure [2]. In an effort to recover the remaining oil, gases such as CO_2 or chemicals such as surfactant and polymer are injected into the reservoir and their properties are exploited to increase the overall sweeping efficiency. This is referred to as tertiary production process. The recovery mechanism using surfactant and polymer, which is the focus of this thesis, works as follow. Polymer increases the viscosity of water making the mobility difference between water and oil favourable to inhibit the growth of viscous fingers. Surfactant, on the other hand, reduces the interfacial tension between water and oil interface increasing the capillary number which is important for oil mobilization. This results in

a production of 30-60% of OOIP. Interested readers may read more about enhanced oil recovery using chemical flooding (cEOR) in the books and review articles [3–8].

1.2   Injection sequence in cEOR
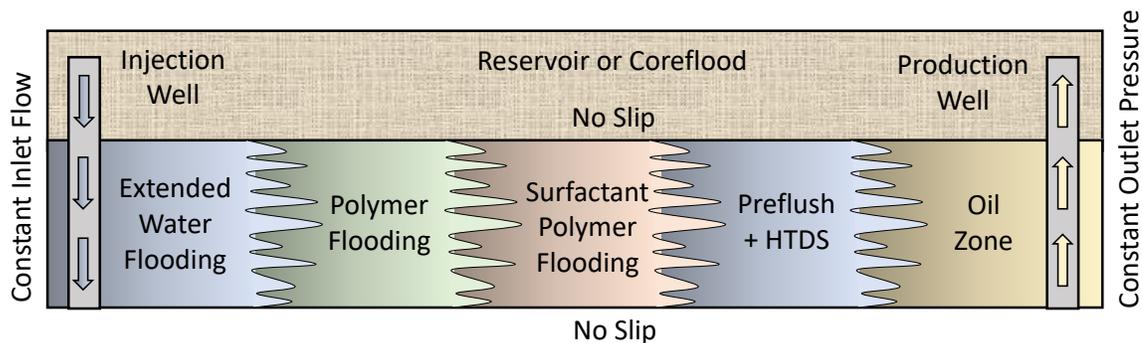
Fig. 1.1.   Schematic of the injection sequence followed in a typical SP flooding process with the boundary conditions used.

The schematic of injection sequence (from left to right) followed in a typical flooding process is shown in figure 1.1. Laboratory scale *coreflood experiments* are often performed to study the performance of the injection sequence as well as the mechanism of oil recovery by surfactant and polymer flooding process. We use Berea sandstone for conducting these coreflood experiments details of which are given in Chapter 2. As seen from figure 1.1, at first, a brine solution (water + HTDS) is injected into the coreflood which results in $\approx 30\%$ recovery of original oil in placed for experiments. At this point, the trapped oil can not be mobilized by simply water flooding and hence EOR must start. Thus, a surfactant-polymer (SP) slug is injected where surfactant and polymer work in synergy to reduce the interfacial tension as well as instabilities at the water-oil interface. This is followed by a polymer slug which helps in maintaining a favourable pressure gradient and high sweeping efficiency by creating a mobility buffer. Finally, chase water (also known as extended water flooding) is injected to flush out the chemicals as well as any trapped oil from the coreflood. In the next

section, we discuss in detail the mechanism of oil recovery by surfactant and polymer flooding.

1.3 Surfactants

Surfactants are amphiphilic in nature meaning they have both hydrophobic (water-loving head) and hydrophilic (water-hating tail) groups. In our work, we use a combination of PETROSTEP® S-13D HA (Alcohol Alkoxy Sulfate) and A6 (Alkyl Benzene Sulfonate) and Huntsman SURFONIC® L series co-solvent at total chemical concentration of 8000 ppm for the SP slug. The structure of surfactant and their orientation in water-oil system is given in figure 1.2. Surfactants are generally classified based on the ionic nature of their head groups or hydrophilic groups [9, 10], critical micelle concentration (CMC) [11] or solubilization ratio.
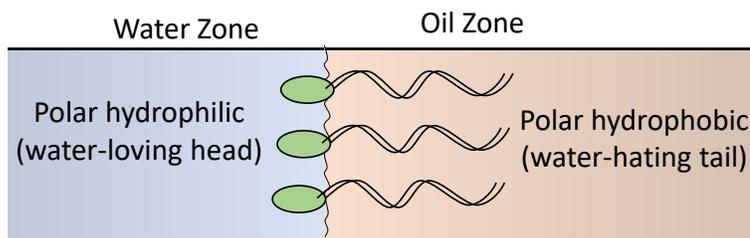
Fig. 1.2. Structure of amphiphilic surfactant molecule with polar hydrophobic head and polar hydrophilic tail and its orientation in oil-water system.

CMC is the critical concentration above which surfactant monomers aggregates to form micelles [12]. Surfactant monomers concentrate at the interface when added to a water-oil system. This releases the system free energy and lowers the interfacial tension (IFT) at the interface [13]. As the surfactant concentration is increased beyond a critical value called CMC, the extra monomers group together to form micelles. Increasing the concentration further results in just the increase of micelles concentration. CMC mainly depends on salinity, temperature and electrolyte concentration

among many other factors. Figure 1.3 show the progress of micelle formation as the concentration is increased along with the reduction in IFT. After CMC, IFT remains constant or increases slightly depending on the type of surfactant.
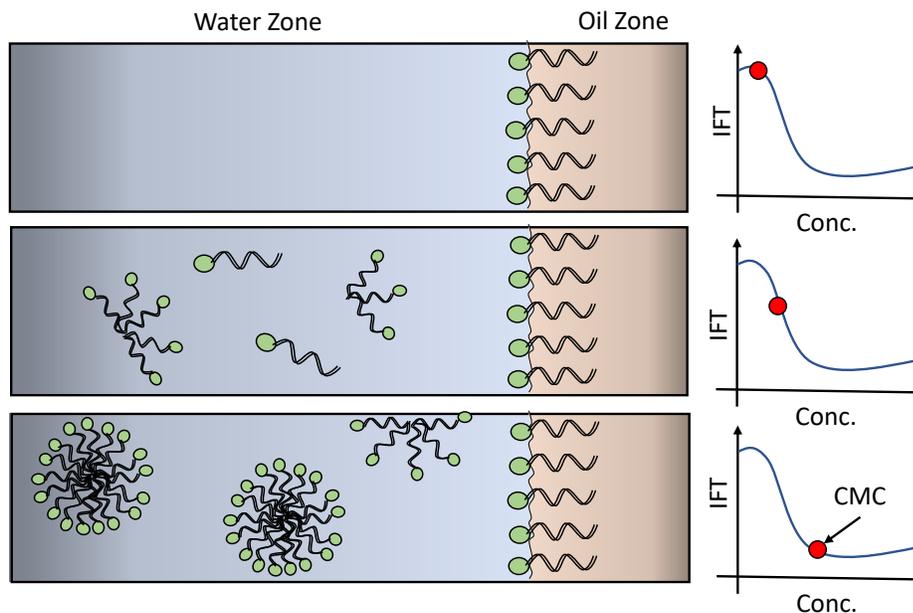
Fig. 1.3. Surfactant molecules saturate at the interface. Increase in surfactant concentration results in reduction of IFT at the interface and micelles start to aggregate. This continuous till the concentration reaches CMC after which the IFT remains almost constant.

1.3.1 Microemulsion systems

Surfactant-water-oil system in the presence of micelles forms microemulsion which is characterized as clear, thermodynamically stable and isotropic in nature [14]. The phase behaviour of microemulsion system depends strongly on the salinity of brine or water phase. At lower salinity, surfactant acts as water soluble aggregating mi-

celles in the water phase, thus creating an oil/water-external system. This system is known as oil-in-water or Winsor I microemulsion system. At higher salinity, surfactant solubilizes in oliec phase and forms a water-in-oil or Winsor II system [15]. At some intermediate salinity range, all the three phases (water-external, oil-external and micro-emulsion) co-exists in the system. The salinity at which water and oil are equally solubilized in microemulsion is called as optimum salinity. In EOR applications, efforts are made to maintain the salinity near the optimum value during the surfactant injection. We will discuss this phase behavior modeling in more detail in section 2.5. Figure 1.4 shows the effect of salinity on microemulsion phase behaviour where monomers are seen to aggregate to form emulsions either in water phase or oil phase depending on the salinity while bi-continuous microemulsion is present at some intermediate salinity range.
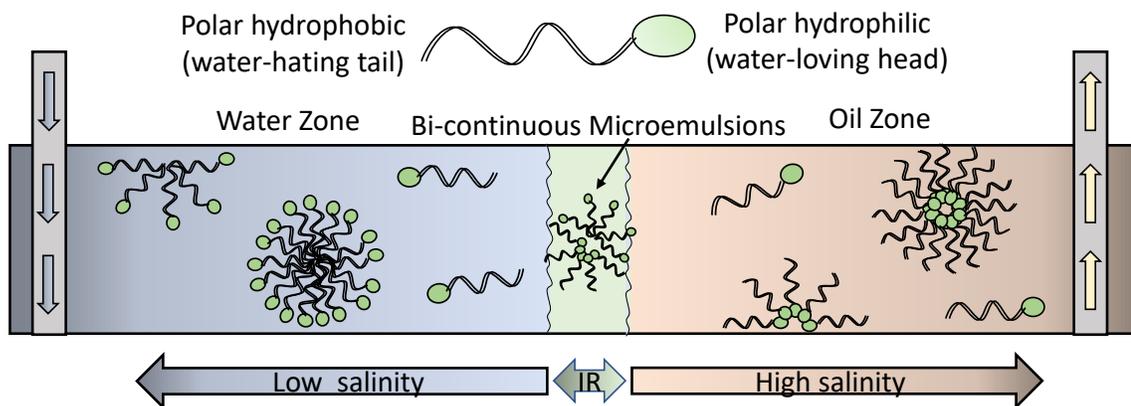
Fig. 1.4. Effect of salinity on microemulsion formation. At lower salinity, surfactant solubilizes in water phase resulting in a Winsor type I system. At some intermediate range (IR) of salinity, three phases exists in the system. At high salinity, surfactant solubilizes in oil phase resulting in a Winsor type II system.

1.3.2 Surfactant flooding

Surfactant helps in the reduction of residual oil saturation either by reducing the IFT of water-oil system or by wettability alteration. During the wettability alteration, oil is driven out of the permeable rock by a spontaneous imbibition of water which shifts the wettability of reservoir from a oil-wet to a water-wet system. A comprehensive review on wettability is given in [16]. It has been observed that the wettability alteration mechanism is more important when IFT is high. Also, it is mainly considered in carbonate cores as they are oil-wet [3, 17]. Since the coreflood under investigation in this work is water-wet Berea sandstone, modeling wettability alteration process is out of the scope of this thesis.

On the other hand, IFT is important with or without the presence of wettability alteration. Reduction in IFT results in lower capillary forces which drives the oil out towards the surface. [18]. To understand the relation between residual oil saturation and IFT, we need to first define Capillary number. Capillary number represents the ratio of viscous drag forces and interfacial tension across the interface of two immiscible fluids. It is a dimensionless quantity given as:

$$Ca = \frac{\mu_w u_w}{\sigma_{ow}} \tag{1.1}$$

where μ_w is viscosity of displacing fluid (water) and u_w is characteristic Darcy velocity of displacing fluid (water). σ_{ow} is IFT between the displacing (water) and the displaced fluid (oil). Experimental data shows that increase in capillary number results in decrease in residual oil saturation [19]. The resulting curve is called Capillary Desaturation Curve (CDC) which is studied in more details in section 2.8.

Capillary number is in the range of $10^{-6} - 10^{-7}$ for a typical water flooding application [9]. Surfactant can help reduce the IFT from 20-30 mN/m to as low as 0.01 mN/m (ultra-low IFT) [20, 21] which translates to an increase in capillary number by 10^3 times. It has been found that residual oil saturation as high as 90% can be achieved by increasing the capillary number to $10^{-3} - 10^{-4}$ [22].

1.4 Polymers

Polymers are long chain molecules that are made by chemically linking together smaller blocks called monomers. In EOR, polymer performance depends on its molecular weight, solvent quality, temperature and the degree of hydrolysis among many other factors [23]. In our study, partially hydrolyzed poly-acrylamide polymer (SNF Flopam 3330) which is commonly known as HPAM is used. This polymer has found a great interest in EOR as it is relatively inexpensive and has good viscosifying properties [24, 25]. The addition of polymer increases the solvent (water) viscosity and creates a mobility buffer that reduces the mobility contrast between the displacing and the displaced fluid. However, polymer viscosity is a function of shear rate, salinity and its concentration. Hence, accurate modeling of polymer flooding performance needs understanding of dependence of each of these factors on the polymer viscosity. We study this in detail in section 2.3.

1.5 Polymer flooding

The stability of an immiscible displacement process is governed by the mobility ratio which is defined as the mobility of displacing or wetting fluid (λ_w) to the non-wetting or displaced fluid (λ_n):

$$M = \frac{\lambda_w}{\lambda_n} = \frac{k_{rw}/\mu_w}{k_{rn}/\mu_n} \tag{1.2}$$

where k_{rw} and k_{rn} is the relative permeability of wetting (water) and non-wetting (oil) phases, respectively. The onset of instabilities at the interface of an immiscible displacement process depends on the mobility ratio. A mobility ratio greater than unity ($M > 1$) in deemed as unfavourable as it causes instability at the interface reducing the volumetric sweeping efficiency [8, 25]. In this case, displacing fluid penetrates or fingers into the displaced fluid resulting in *viscous fingering* [26]. Viscous fingering is undesirable in EOR applications because it promotes itself and reduces the overall production by early breakthrough. It is clear from Eq. (1.2) that one way to reduce

the mobility ratio and thus prevent the growth of viscous fingers is to increase the viscosity of displacing fluid ($M < 1$). Hence, polymer solutions are added to the displacing fluid to promote stable displacement of the interface. Figure 1.5 shows the mechanism of growth of viscous fingering and its prevention by polymer flooding.

The mixing in reservoir or coreflood is a result of heterogeneity of the medium as well as physical dispersion. For highly heterogeneous cores, dispersion is dominated by heterogeneity. However, when core is modeled as homogeneous, dispersivity must be considered to model the right degree of mixing [27]. Understanding viscous fingering and fluid mixing is important as in later chapters, for 1D corefloods, we will introduce artificial dispersivity to take into account the heterogeneous nature of the core. Here, we simulate two cases in section 1.8: with polymer and without polymer and show that the polymer flooding can be used to prevent the onset of instabilities and hence growth of viscous fingering.
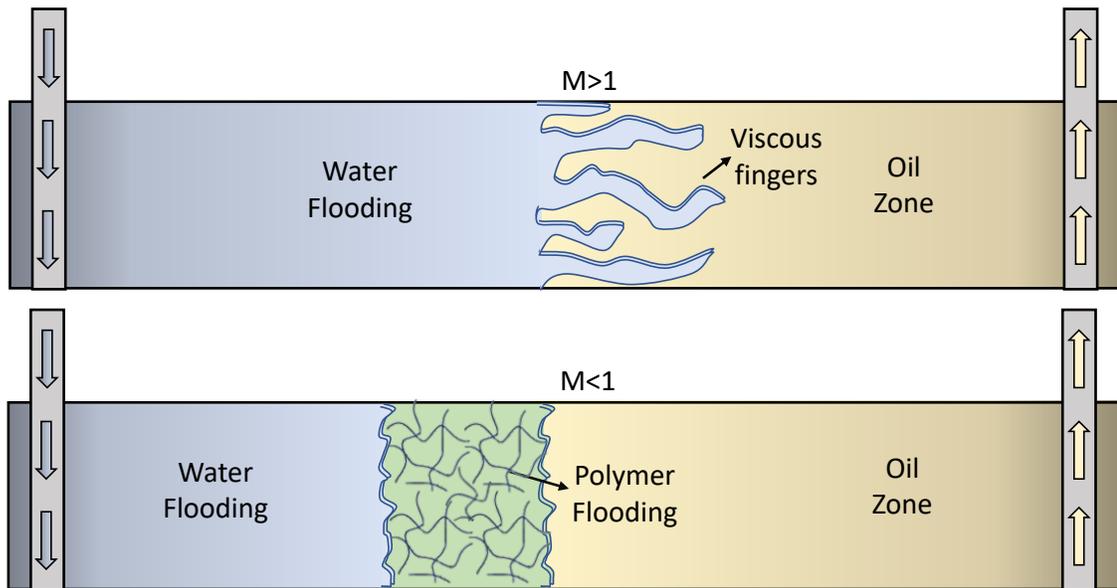
Fig. 1.5. Mobility ratio (M) > 1 results in instability at the interface and onset of viscous fingers. Polymer flooding inhibits the growth of viscous fingering by reducing the mobility of displacing fluid (M<1).

1.6 Governing Equations

Our aim is to study the immiscible displacement of non-wetting phase by a wetting phase in a porous media in the presence of solvents. In this thesis, we use UTCHEM simulator, originally developed by Pope and Nelson [27], to model this multicomponent two-phase imbibition process in porous media. The components (κ) that are present are wetting phase (water), non-wetting phase (oil), and solvents (surfactant and polymer). We start with the transport equation (mass balance equation) for the wetting phase (neglecting adsorption) [28]:

$$\phi\frac{\partial(S_w C_{\kappa w})}{\partial t} + \nabla \cdot (\mathbf{u}_w C_{\kappa w} - \phi S_w \mathbf{D}\nabla C_{\kappa w}) = R_\kappa, \quad \kappa = w, s, p \tag{1.3}$$

where ϕ is the medium porosity, $\mathbf{u}_w$ is the Darcy velocity of the wetting phase, $C_{\kappa w}$ is the concentration of each component in the wetting phase (i.e. pure wetting phase (w), surfactant (s), and polymer (p)) and S_w is the saturation of wetting phase. The total saturation of wetting and non-wetting phase is always constrained as $S_w + S_n = 1$. The phase flux $\mathbf{u}_l$ for wetting ($l = w$) and non-wetting ($l = n$) phases with the relative permeability k_{rl}, viscosity μ_l, and density ρ_l can be calculated from the Darcy's law as:

$$\mathbf{u}_l = \frac{-k_{rl}}{\mu_l}\mathbf{K}(\nabla p_l + \rho_l g\nabla z), \quad l = w, n \tag{1.4}$$

where $\mathbf{K}$ is the absolute permeability tensor of medium. The gravity, pressure and vertical depth are denoted by g, p and z respectively. We neglect any gravitational effects. The phase pressure difference wetting p_w and non-wetting phase p_n is called capillary pressure p_c which is a function of saturation S_w and is given as:

$$p_c(S_w) = p_w - p_n \tag{1.5}$$

The relative permeability of each phase k_{rl} is modeled using Brooks and Corey equation as:

$$k_{rl} = k_{rl}^0(\overline{S}_l)^{n_l} \tag{1.6}$$

where n_l is the Corey exponent for l^{th} phase and

$$\overline{S}_l = \frac{S_l - S_{lr}}{1 - \sum\limits_{l=w,n} S_{lr}} \tag{1.7}$$

is the normalized saturation. More details are given in section 2.8. With such a model, p_c is given as:

$$p_c(S_w) = p_{c,0}(\overline{S}_l)^{\frac{-1}{\lambda_b}} \tag{1.8}$$

where λ_b is a fitting parameter which depends on pore size distribution and $p_{c,o}$ is a model fitting parameter. The dispersion tensor $\mathbf{D}$ in Eq. (1.3) with molecular diffusion coefficient D_m is given as [29]:

$$D_{ij} = (D_m + (\alpha_T|\mathbf{u}_w|))\delta_{ij} + (\alpha_L - \alpha_T)\frac{u_{wi}u_{wj}}{|\mathbf{u}_w|} \quad i = x, y \tag{1.9}$$

where δ_{ij} is the Kronecker delta function, α_L and α_T are the longitudinal and transverse dispersivities of solvents in the wetting phase, respectively. Finally, R_κ in Eq. (1.3) accounts for the injection or production of each component through injection and production wells similar to a source or sink term. The well model used is based on [30,31]. We maintain a constant flow rate from the injection well while the pressure well is kept at a constant pressure. A no flow boundary condition is assumed for all the boundaries (i.e. no flux through the transverse boundaries) (refer to figure 1.1).

We use the Implicit Pressure Explicit Saturation (IMPES) scheme [32–34] to solve for the unknown pressure field p_n and saturation field S_w. Recall that the solvents are only soluble in the wetting phase, therefore the non-wetting phase saturation S_n can be calculated as $S_n = 1 - S_w$, once S_w is known. At first, the transport equation over all the volume occupying components is summed together to derive the pressure equation. The resulting linear system of equations is then implicitly solved using the known saturation at the current time step to determine the pressure field. This pressure field is substituted in Darcy's equation (refer to Eq. (1.4)) and using the capillary pressure relation (refer to Eq. (1.5)), the phase velocity field at the current time step is derived. The velocity field is then used in Eq. (1.3) to explicitly solve for

saturation at the next time step. A more detailed explanation on IMPES scheme is available in [32, 35, 36]. It is clear that in IMPES scheme, the pressure and saturation equations are decoupled by an explicit treatment of capillary pressure equation. In heterogeneous media, the capillary forces may change the saturation field [32] and hence very small time steps should be used. In all the problems studied in this thesis, we maintain a maximum CFL of 0.01 using appropriate time steps [28]. Next, we present the validation of IMPES with the analytical solution of Buckley-Leverett equation where Total Variation Diminishing (TVD) scheme is used to capture the sharper shocks with minimum numerical dispersion [37].

1.7 Validation with Buckley-Leverett equation

The saturation profile obtained by solving Eq. (1.3) in one-dimentional is a well-known Buckley-Leverett profile which features a piston-like displacement of shock front. Buckley and Leverett (BL) in 1942 proposed an equation to model immiscible displacement process of oil recovery (non-wetting phase) by water flooding (wetting phase) in one dimensional reservoir [38]. BL equation provides an analytical solution to capture the advance of wetting and non-wetting phases displacement front using conservation of mass for each phase under the assumption of incompressible and immiscible flow, and neglecting the gravity and capillary forces (without dispersion):

$$\frac{\partial S_w}{\partial t} + \frac{\partial}{\partial x}\left(\frac{Q}{\phi A} f_w(S_w)\right) = 0, \tag{1.10}$$

where ϕ is porosity of the reservoir, $S_w(x,t)$ is the wetting phase saturation with $S_w(x,0) = S_0(x)$, Q is the total flow rate through the cross-sectional area A of the reservoir. $f_w(S_w)$ is called the fractional flow function of wetting phase and is given by:

$$f_w(S_w) = \frac{\lambda_w}{\lambda_w + \lambda_{nw}} = \frac{\dfrac{k_{rw}}{\mu_w}}{\dfrac{k_{rw}}{\mu_w} + \dfrac{k_{rn}}{\mu_n}} \tag{1.11}$$

where λ_l is the mobility of the lth phase as defined in Eq. (1.2) with $l = w$ (wetting) and $l = n$ (non-wetting). The fractional flow curve given in Eq. (1.11) has a single inflection point and can be used to determine the speed of the shock by constructing a tangent through it in a particular fashion as first discussed in [39].

We use Corey's equation to model relative permeability of each phase (refer to Eq. (1.6)). Relative permeability curves with end point relative permeability $k_{rl}^0 = 1.0$, Corey exponent $n_l = 2.0$ and residual saturation $S_{lr} = 0.0$ for both the phases are given in figure 1.6(a). Here, we validate the simulation model with the analytical BL equation for three cases of different relative viscosity ratios $\mu_r = \dfrac{\mu_n}{\mu_w} = 0.5, 1.0, 2.0$. Figure 1.6(b) shows the variation of fractional flow (water) with its saturation for the three viscosity ratios considered here.

(a) Relative permeability curve

(b) Fractional flow curve for water phase
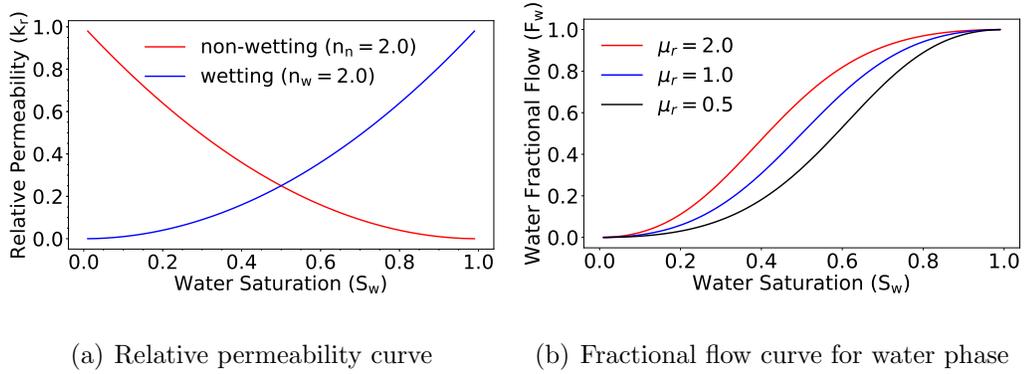
Fig. 1.6. Left: Relative permeability curve modeled using Corey's equation with the exponents for the wetting $n_w = 2$ and non-wetting phase $n_n = 2$. Right: Fractional flow curve of the wetting phase for three different relative viscosity ratios ($\mu_r = 0.5, 1.0, 2.0$)

Figure 1.7 shows that the simulation model with TVD flux-limited higher-order scheme accurately captures the onset of shock across the interface. The model can be seen to accurately predict the advancement of shock for all the three viscosity ratios (the results are plotted after one day).
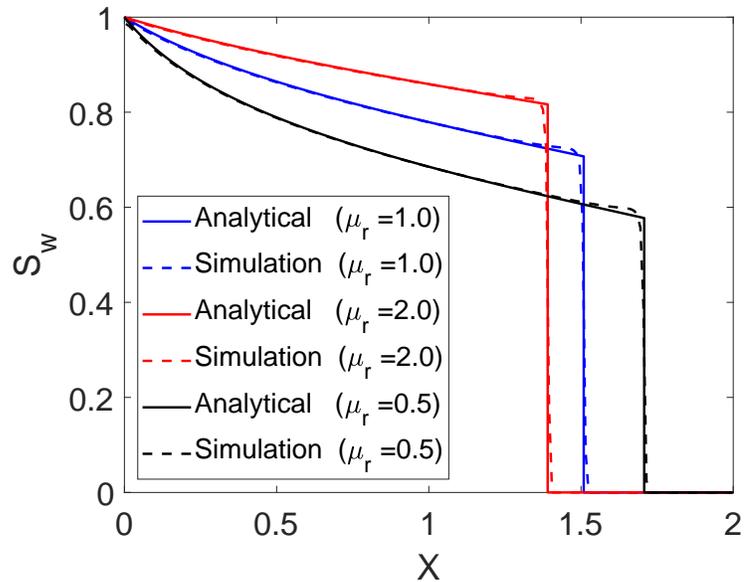
Fig. 1.7. Validation of simulation model with analytical Buckley-Leverett equation for different values of viscosity ratios. The grid size used is 200 x 1 and the results are plotted after one day.

1.8 Case study: Viscous fingering

In this section, we simulate two cases of water flooding in 2D coreflood 1) with no polymer 2) with polymer (0.25% wt.) to show the mechanism of viscous finger suppression by polymer injection. The core is considered to be heterogeneous with a constant porosity ($\phi = 0.21$) and a constant initial water saturation ($S_w = 30\%$). The permeability field is generated using Kriging process where log-k variance of 0.1 and covariance lengths of 0.1 times the domain size are used [28]. The domain is discretized into 400 x 80 grid blocks. The relative permeability curves used are given in figure 1.8 where at low Ca, the end point relative permeability and Corey exponent for Eq. (1.6) are obtained from water flooding Berea sandstone. Relative permeability curves at high Ca are assumed to be linear with zero residual water and oil saturation [40]. The shape of relative permeability curves (i.e. k_{rw}, k_{rn}) greatly affects the mobility ratio [41]. It has been shown experimentally [42] as well

as numerically [43] that the mobility ratio across the shock M_s should be considered while assessing the onset of instability. In such studies, the shock mobility ratio is determined from the fractional flow curve. However, in case of surfactant-polymer flooding, it is not feasible to construct fractional flow curve as there is a spatio-temporal change in relative permeability curves with an increase of capillary number. Therefore, we use end-point relative permeability (i.e. k_{rw}^0, k_{rn}^0) to assess and predict the onset of instability [28, 44].
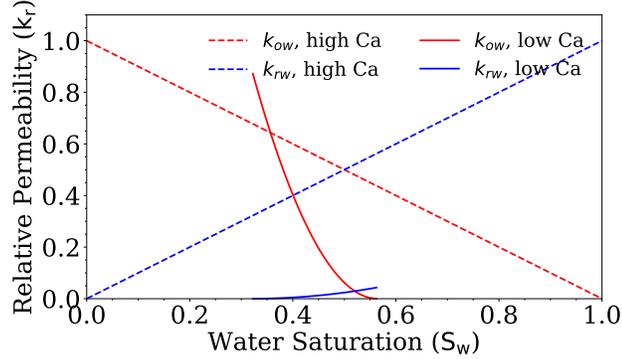
Fig. 1.8. Relative permeability curves at low Ca and high Ca used for the case study of polymer injection.

We use relative viscosity ratio $\mu_r = \mu_o/\mu_w$ 20.0 for both of these cases. Polymer viscosity is assumed to be a function of concentration only and the salinity or shear effects on polymer viscosity are neglected. Eq. (2.1) was used to model polymer viscosity where the fitting parameters were obtained from the rheological tests conducted on the polymer at different concentrations (refer to section 2.3). Figure 1.9 shows polymer flooding in action. For Case 1. with no polymer, the mobility ratio calculated using the end point relative permeabilities is $M = 2.5$ (> 1). The viscous fingers can be clearly seen to be penetrating through the oil phase. For Case 2. addition of polymer (0.25% wt.) reduces the mobility ratio to $M = 0.5$ (< 1) which inhibits the growth of this viscous fingers. A piston-type displacement can be observed for this case. An important observation is that for the case with no polymer,

the water can be seen to breakthrough earlier than the case with polymer which is deemed unfavourable as this results in reduced overall oil recovery.

T=0.04 PV   T=0.04 PV

T=0.08 PV   T=0.08 PV

T=0.12 PV   T=0.12 PV

(a) Case 1. No polymer   (b) Case 2. 0.25% wt polymer
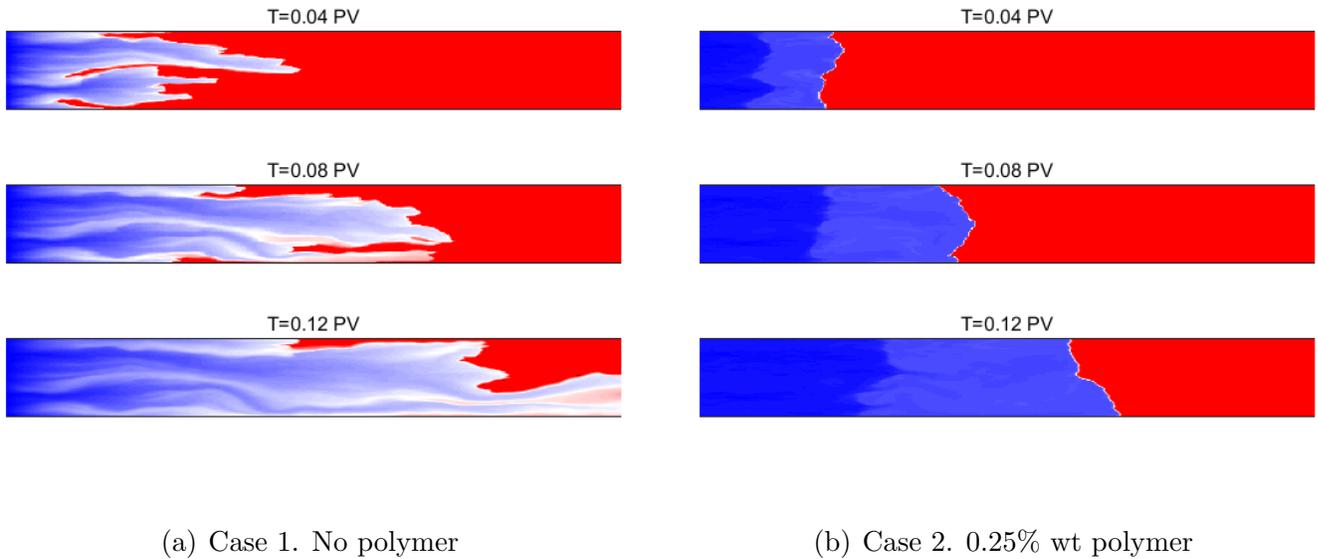
Fig. 1.9. Polymer flooding in action. On left: water flooding with no polymer results in viscous fingering and hence early breakthrough. On right: water flooding with polymer (0.25%) creates a mobility buffer and results in a piston-like displacement.

1.9 Conclusion

This chapter presented a brief introduction to the different technologies available for mobilization of trapped oil in reservoirs and the need for chemical EOR process. Then we studied the injection sequence employed for a typical surfactant-polymer flooding scenario. We also discussed the mechanism of oil recovery by surfactant and polymer flooding. Surfactant flooding reduces the IFT and helps in easy mobilization of trapped oil. Polymer flooding creates a mobility buffer between the oil and water phase making the mobility ratio favourable. This prevents the growth of viscous fingers and increases the sweeping efficiency. We also simulated this mechanism of

viscous finger suppression for two cases of water flooding, with and without polymer, in a 2D heterogeneous core. We presented a mathematical model for immiscible displacement of two phase flow through a porous media. Finally, we validated the simulation model with the benchmark one-dimensional Buckley-Leverett equation. In the next chapter, we will discuss various physical models for surfactant and polymer that govern the mechanism of oil recovery and the experiments that are carried out to calibrate these models.

2. PRIOR MODEL CALIBRATION USING COREFLOOD EXPERIMENTS

A version of this chapter has been published as:

Soroush Aramideh, Rituraj Borgohain, **Pratik Naik**, Cliff T. Johnston, Pavlos P. Vlachos, and Arezoo M. Ardekani. "Multi-objective history matching of surfactant-polymer flooding." *Fuel* 228 (2018): 418-428

2.1 Abstract

The problem of interest here is a process called surfactant-polymer (SP) flooding [45]. SP flooding is one the most promising enhanced oil recovery technique and involves injection of surfactant and polymer into an oil reservoir to boost the production [5, 46]. Polymer is injected to improve the sweeping efficiency and hinder flow instabilities [47] while surfactant is responsible for mobilizing the trapped oil in small pores by reducing the interfacial tension between aqueous and oleic phases [48, 49]. Despite the elegant oil recovery mechanism in SP flooding and its high efficacy in the controlled laboratory experiments, it has shown poor performance in field experiments [50, 51]. This is due to intrinsic complexity of SP flooding and large uncertainty present in this process [52]. To design a successful SP flood, it is necessary to have a reliable and accurate SP flood model. The first step towards constructing a reliable SP flood model that can be used for design and optimization purposes is *model calibration* which is the focus of this chapter.

Accurately capturing this elegant mechanism of oil recovery by both surfactant and polymer injection requires knowledge of various physical models imitating these processes [53]. This can be obtained by performing separate experiments that address each of these physical processes individually. These are referred to as *support-*

experiments in this thesis. In this chapter, we discuss the details of these laboratory scale experiments that were used to build a mechanistic model of SP flooding process. We use these support-experiments in an initial calibration stage with an ultimate objective to history match surfactant-polymer flooding in Berea coreflood (BCF) experiments, details of which are given in section 2.12. We model the surfactant-polymer flooding process in corefloods using reservoir simulator UTCHEM v9.82 [27].

2.2   Support experiments for initial model calibration

In this section, we discuss different support-experiments that were preformed to calibrate various physical models of SP flooding. All the experiments are carried out on Berea sandstone at room temperature. Crude oil was obtained from the reservoir field. Table 2.1 shows the properties of oil and synthetic brine solution used in the experiments.

Table 2.1.

Properties of oil and brine used in the experiments.

Property	Oil	Brine
Viscosity (cP)	13.8	1.0
IFT (dyne/cm)	22	22
Anions (meq/ml)	-	0.145
Divalent cations (meq/ml)	-	0.0097

We develop the SP flood model by dividing the entire process of SP flooding in three separate sub-processes 1) surfactant flooding 2) polymer flooding and 3) displacement process. The physical models describing each of these processes are calibrated individually by using different calibration methods as shown in figure 2.1. The main purpose of performing these different experiments is to reduce the uncertainty in the prior knowledge of the model parameters by using laboratory measured

data. This is crucial as the parameter space of the different physical models (such as polymer viscosity, surfactant phase behavior and IFT, multiphase displacement models etc.) governing the SP flood process is of high dimensions which often results in a non-unique solution [54]. Therefore, performing separate experiments help us to gain a better prior understanding of each of these models. The remainder of the chapter briefly explains these support experiments.
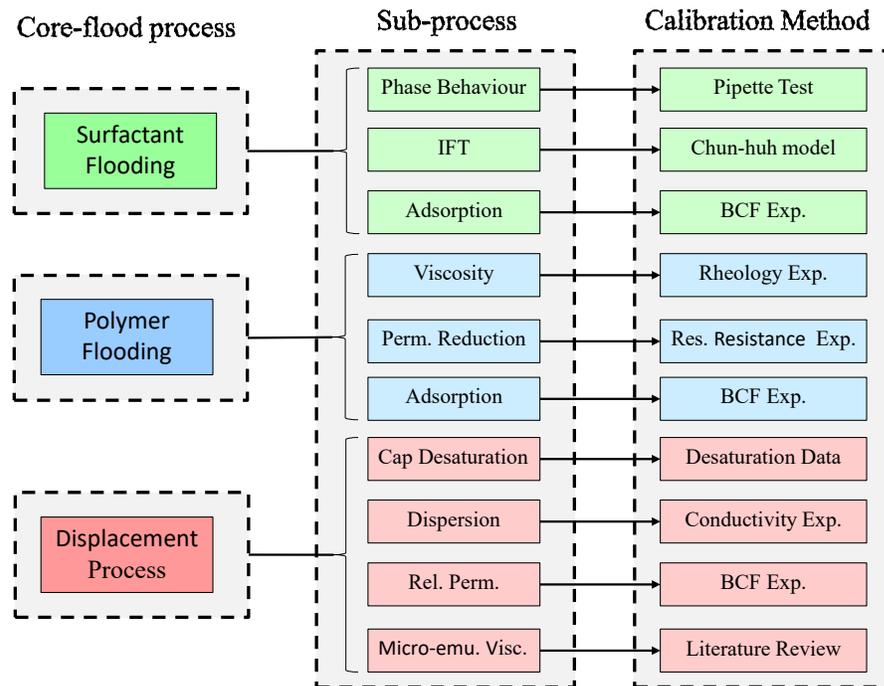
Fig. 2.1. Classification of SP flooding process as surfactant flooding, polymer flooding and displacement process. The sub-processes characterizing each process is also shown with the adapted calibration method.

2.3 Modeling polymer flooding processes

The viscosity of a polymer solution depends on its concentration, salinity and shear rate [27]. Therefore, each of these factors should be considered when model-

ing polymer behavior in corefloods. We perform a series of rheology experiments on HPAM 3230 and 3330 to quantify polymer dependence on these factors. The variation of polymer viscosity with its concentration and salinity is modeled using Flory Huggins equation [55]:

$$\mu_p^0 = \mu_w(1 + (A_1 + A_2 C_p^2 + A_3 C p^3)C_{SEP}^{S_p}), \tag{2.1}$$

where μ_p^0 is the viscosity of polymer when shear rate is zero, C_{SEP} is the effective salinity, A_1, A_2, A_3 are the model parameters and S_p is the exponent used in Eq. (2.2).

$$\frac{(\mu_p^0 - \mu_w)}{\mu_w} \propto C_{SEP}^{S_p}. \tag{2.2}$$

Effective salinity C_{SEP} is modeled as:

$$C_{SEP} = \frac{C_{51} - (\beta_p - 1)C_{61}}{C_{11}} \tag{2.3}$$

where C_{11}, C_{51}, C_{61} are water, anion and divalent cation concentrations in the aqueous phase. β_p models the impact of divalent cations on the polymer viscosity [27]. Finally, the shear-thinning polymer viscosity is modeled via Eq. (2.4).

$$\mu_p = \mu_w + \frac{\mu_p^0 - \mu_w}{1 + \left(\frac{\dot{\gamma}_1}{\dot{\gamma}_{1/2}}\right)^{P^\alpha - 1}}. \tag{2.4}$$

where $\dot{\gamma}_{1/2}$ and P^α are fitting parameters. P^α models the curvature of the curve at high shear rate. Figure 2.2 shows the model fitting for polymer viscosity where figure 2.2(a) shows the variation of polymer viscosity with the concentration given by the corresponding model in Eq. (2.1) at 9400-22000 ppm salinity range. Figure 2.2(b) shows polymer viscosity as a function of C_{SEP} (in meq/ml) described by Eq. (2.2). The exponent S_p is the fitting parameter. Figure 2.2(c) shows polymer viscosity as a function of shear rate which is modeled using Eq. (2.4). Polymer and water viscosity used here are measured at the same temperature.
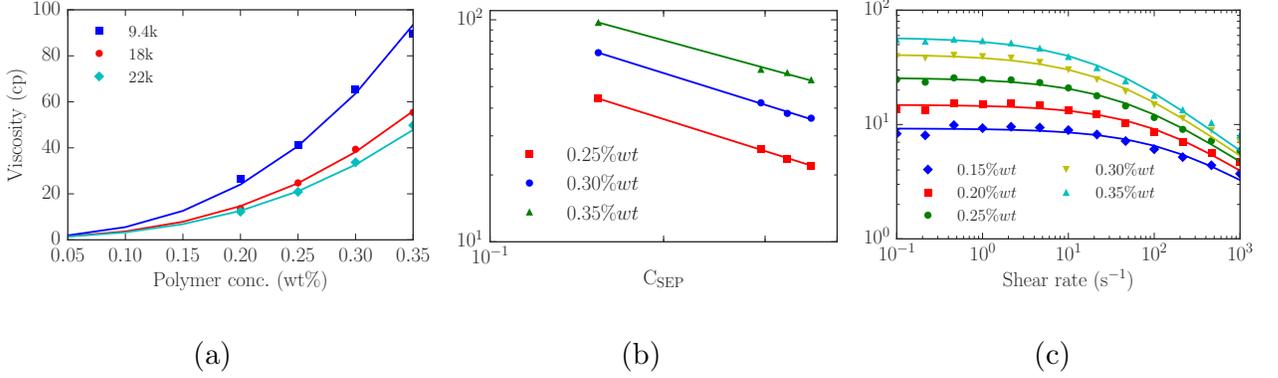
Fig. 2.2. Polymer viscosity is modeled as function of a) concentration using Eq. (2.1) at different salinities b) effective salinity C_{SEP} using Eq. (2.2) at different concentrations and c) shear rate using Eq. (2.4) at different concentrations.

We modify the existing model in UTCHEM to make $\dot{\gamma}_{1/2}$, a function of polymer concentration and salinity (refer to Eq. (2.5)).

$$\dot{\gamma}_{1/2} = (a_1 C_{SEP} + a_0)exp(a_2 C_p), \tag{2.5}$$

where a_0, a_1 and a_2 are fitting parameters. Figure 2.3 shows variation of $\gamma_{1/2}$ for different salinity values which appear to be an exponential function of concentration for HPAM 3330. The corresponding fit is obtained using Eq. (2.5).

2.4 Residual resistance factor (RRF) experiments

Polymer flooding results in mobility reduction of displacing fluid as well as reduction of permeability of porous medium. To model this permeability reduction due to polymer flooding, we perform a separate single phase coreflood experiment details of which are given in table 3.2. The core is initially flooded with water and the absolute permeability is measured. This is then followed by a polymer slug till the pressure reaches a steady state after which the core is again flooded with water and the absolute permeability is measured.
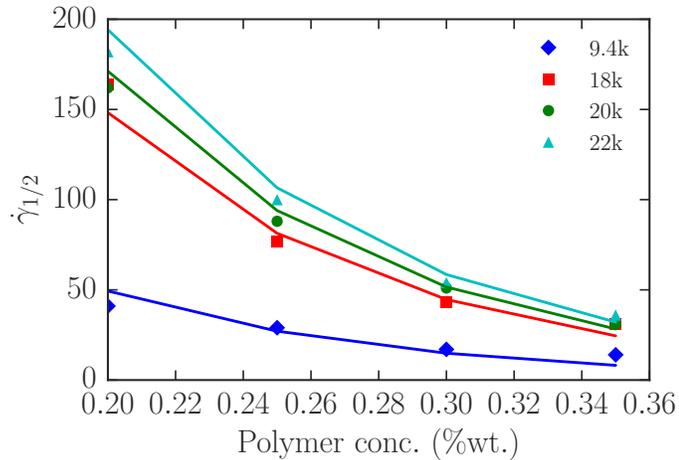
Fig. 2.3. Variation of $\dot\gamma_{1/2}$ with polymer (HPAM 3330) concentration and salinity and the corresponding model fit with Eq 2.5. Salinity is in ppm.

Table 2.2.

Details of the support experiment carried out to calibrate permeability reduction.

Quantity	Value
Length (inch)	6
Diameter (inch)	2
Flow rate (cc/min)	0.25

The permeability reduction factor (R_k) can be calculated as the ratio of effective permeability of water to the effective permeability of polymer. The combined effect of increase in viscosity as well as the reduction in permeability results in the decrease in mobility which is termed as residual resistance (R_F) and modeled as:

$$R_F = R_k \frac{\mu_w}{\mu_p}, \tag{2.6}$$

where permeability reduction factor is given by Eq. (2.7)

$$R_k = 1 + \frac{(R_{k,max} - 1)b_{rk}C_{4l}}{1 + b_{rk}C_{4l}}, \tag{2.7}$$
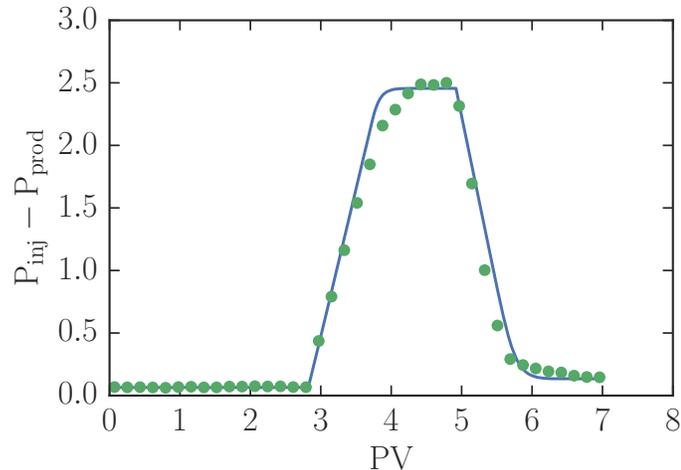
Fig. 2.4. Validation of pressure profile for permeability reduction experiment. Water flooding is followed by a polymer slug till the pressure reaches a steady state value after which core is again flooded with water.

where $R_{k,max}$ is modeled with Eq. (2.8)

$$R_{k,max} = max\left[\left[1 - \frac{c_{rk}(A_1 C_{SEP}^{S_p})^{1/3}}{\left(\frac{\sqrt{k_x k_y}}{\phi}\right)^{1/2}}\right]^{-4}, 10\right], \qquad (2.8)$$

where c_{rk} and b_{rk} are input fitting parameters and k_i is the permeability in i direction. This model assumes that the permeability reduction is irreversible i.e. permeability reduction does not decrease with the decrease in polymer concentration. Figure 2.4 compares the pressure profile predicted by the model (refer to Eq. (2.7)) with the experimental data.

2.5 Phase Behaviour Modeling

To study the phase behaviour of surfactant in water-oil system, we perform phase behaviour equilibrium tests called pipette tests. In these tests, the surfactant solution is added to water-oil system in a test tube at varying salinities and is allowed to stagnate which results in emulsification. Then solubilization ratios are calculated for

both water and oil by visual inspections. We perform these experiments with water to oil ratio 1 and with 0.678 % surfactant concentration. It is worth noting that these tests are done in static conditions and do not probe in-situ mixing [45, 56].

Salinity has a strong influence on surfactant solution phase behaviour (refer to figure 2.5). Depending on the salinity, the surfactant monomers aggregate to form micelles either in the aqueous phase or oleic phase. Generally, the surfactant phase behaviour can be categorized into three types depending on the solubilization of microemulsion in aqueous phase or oleic phase, which are listed below.

2.5.1 Types of microemulsions

1. **Lower Phase microemulsion (Type I):** At low salinity, surfactant is primarily solubilized in aqueous phase and so the oleic phase is relatively surfactant free. Since, the aqueous phase is denser than the oleic phase (the aqueous phase settles at the bottom of the pipette tube as shown in figure 2.5) and microemulsion resides in the lower aqueous phase, it is called lower phase microemulsion. It is also referred to as Winsor Type I or Type II (-) microemulsion.

2. **Upper Phase Microemulsion (Type II):** At higher salinity, the surfactant exhibits completely opposite characteristics as compared with low salinity. The surfactant solubilizes in oleic phase forming two phases: oil-external microemulsion and an excess water phase. This is referred to as Winsor Type II or Type II (+) microemulsion.

3. **Middle Phase Microemulsion (Type III):** There exists all the three phases i.e. excess oil, microemulsion, and excess water at some intermediate salinity range. In this system, the microemulsion is solubilized both in the oleic phase as well as aqueous phase and forms a separate phase in the between water and oil. This is referred to as Winsor Type III or simply Type III. From figure 2.5, it is evident that the middle phase microemulsion onset at around 17.5k ppm salinity

We model the phase behaviour using tie lines, binodal curve and Hands Rule, the details of which can be found in [27, 57, 58]. It is worth noting that the Hands Rule is not able to capture the phase behaviour at higher salinity (refer to figure 2.5). But this should not affect the modeling of BCF experiments in Chapter 3 as all the coreflood experiments are conducted at lower salinity ($\leq$ 17500 ppm).

(a)   (b)
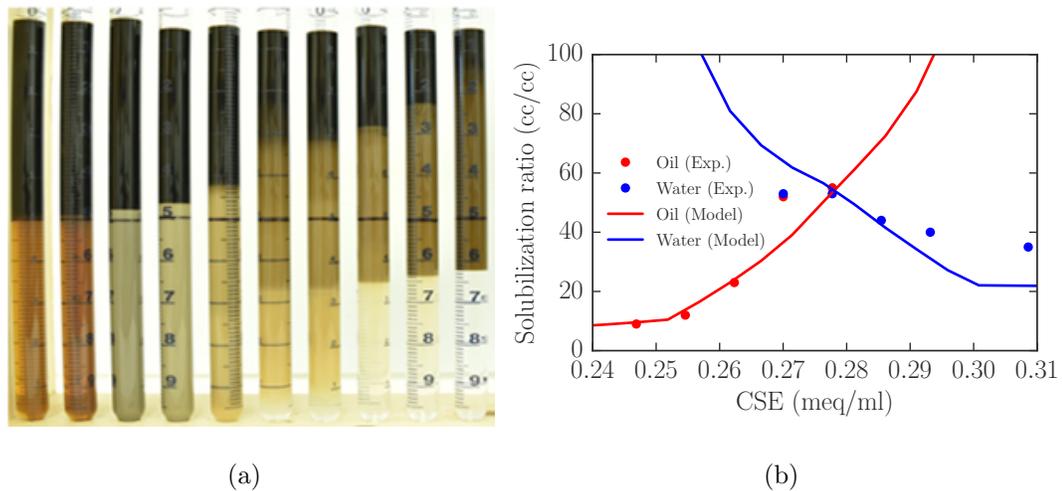
Fig. 2.5. a) Pipette tubes experimental data for salinities ranging from RockHill (RH) to 20k ppm for S13D surfactant for 0.678% concentration. b) Model calibration to the solubilization ratio using Hands Rule.

2.6   Interfacial tension model

We use Chun-Huh equation to model the relation between the interfacial tension (IFT) and the solubilization ratio as:

$$\sigma_{l3} = \sigma_{ow} \exp(-aR_{l3}) + \frac{cF_l}{R_{l3}^2}(1 - \exp(-aR_{l3}^3)), \ l = 1, 2 \tag{2.9}$$

where σ_{l3} is the interfacial tensions between the microemulsion phase and phase l. σ_{ow} denotes the interfacial tension between oil and water in the absence of surfactant. R_{l3}

is the solubilization ratio for l^{th} phase, a and c are fitting parameters. The Hirasaki's correction factor, F_l, is defined as :

$$F_l = \frac{1 - \exp(-\sqrt{con_l})}{1 - \exp(-\sqrt{2})} \qquad (2.10)$$

where $con_l = \sum_{k=1}^{3}(C_{kl} - C_{k3})^2$ and C_{kl} denotes concentration of species k in phase l. Figure 2.7 shows the fit for Chun-huh model with the fitting parameters $c = 0.3$ and $a = 10$.
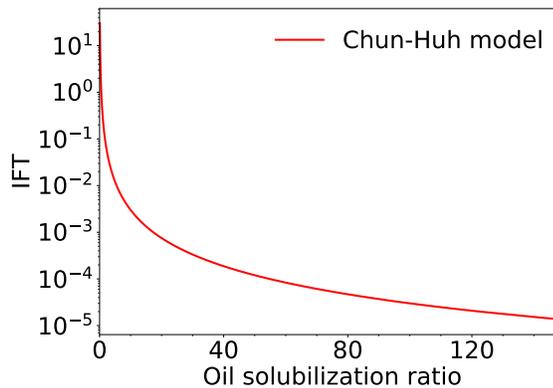
Fig. 2.6. IFT modeled using Chun-huh relation with the fitting parameters: $c = 0.3$ and $a = 10$.

2.7 Microemulsion viscosity

The microemulsion viscosity is modeled using the following model:

$$\mu_3 = C_{13}\mu_w \exp[\alpha_{V1}(C_{23} + C_{33})] + C_{23}\mu_o \exp[\alpha_{V2}(C_{13} + C_{33})]$$
$$+ C_{33}\alpha_{V3} \exp[\alpha_{V4}C_{13} + \alpha_{V5}C_{33}] \qquad (2.11)$$

where $\alpha_{V1}, \alpha_{V2}, \alpha_{V3}, \alpha_{V4}, \alpha_{V5}$ are fitting parameters. μ_w and μ_o are water and oil viscosity and C_{kl} is concentration of kth component in l^{th} phase where $k = 1$: water, $k = 2$: oil, $k = 3$: surfactant, and $l = 1$: aqueous, $l = 2$: oleic, and $l = 3$: microemulsion. Figure 2.7 shows microemulsion viscosity variation with the oil solubilization

ratio obtained using Eq. (2.11) where experimental stabilization ratios were used to determine the microemulsion viscosity.
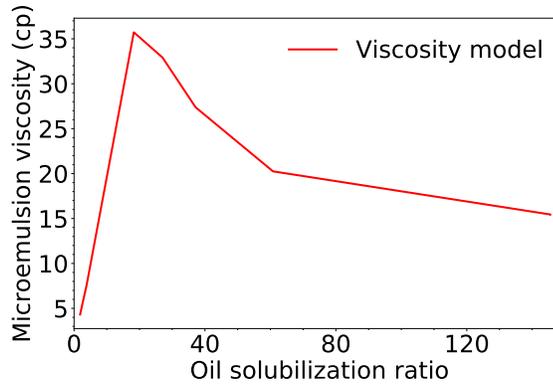
Fig. 2.7. Microemulsion viscosity calculated using microemulsion viscosity model with fitting parameters $\alpha_{V1} = 1.0$, $\alpha_{V2} = 3.0$, $\alpha_{V3} = 0.0$, $\alpha_{V4} = 0.9$, $\alpha_{V5} = 0.7$.

2.8 Capillary desaturaton

Two-phase immiscible displacement at pore scale level is governed by dimension-less capillary number Ca (refer to Eq. (1.1)). Capillary desaturation curve (CDC) relates the residual oil saturation with Ca and describes the phase entrapment at a given Ca [59]. To determine the CDC, we conduct a water-flooding experiment on a Berea core of high permeability (Length: 6", Diameter: 2"). The details of experimental setup for the initial state is given Table 2.3.

The flow rate is controlled to change the capillary number Ca. Flow rate is varied from 0.09 to 100 cc/min which resulted in S_{or} as low as $\approx 10\%$. This translates to a reduction of 0.002 - 0.003 mN/m in IFT as shown in the Figure 2.8. This agrees well with the literature [60].

Table 2.3.
Details of the initial experimental setup for the capillary desaturation experiment.

Property	Value
Flow rate (cc/min)	0.09
Porosity	0.23
Permeability(mD)	700
Viscosity (cP)	2.14, 14.2
IFT (mN/m)	2.80
Capillary number	2.5E-06
Core age	2 weeks
Co-solvent	20% Vol/Vol
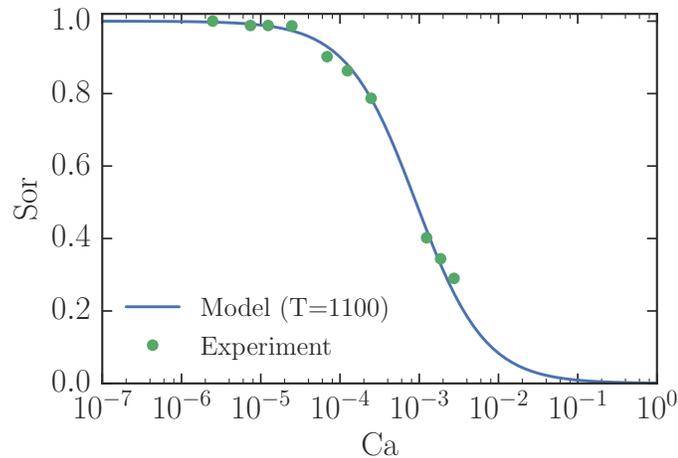
Fig. 2.8. Capillary desaturation experimental results and its corresponding model fit. An increase in capillary number can be seen to reduce the residual oil saturation.

The residual saturation of phase l (S_{lr}) in figure 2.8 is modeled based on the capillary number as shown below:

$$S_{lr} = min\left(S_l, S_{lr}^{high} + \frac{S_{lr}^{low} - S_{lr}^{high}}{1 + T_l(Ca_l)^{\tau l}}\right) \qquad (2.12)$$

where S_l is the saturation of phase l. τ_l and T_l are fitting parameters for trapped saturation data for phase l ($l = 1$: aqueous, $l = 2$: oleic, $l = 3$: microemulsion). S_{lr}^{high} and S_{lr}^{low} are trapped saturation at high and low capillary numbers, respectively. We use a value of one for τ_l. The endpoint relative permeability for each phase increases in a very predictable way as the trapping number increases. The following equation is used to modify the endpoint relative permeability based on residual saturation of conjugate phase:

$$k_{rl}^0 = k_{rl}^{0,low} + \frac{S_{l'r}^{low} - S_{l'r}}{S_{l'r}^{low} - S_{l'r}^{high}} \left(k_{rl}^{0,high} - k_{rl}^{0,low}\right) \tag{2.13}$$

where $S_{l'r}$ is the residual saturation of the conjugate phase. $k_{rl}^{0,high}$ and $k_{rl}^{0,low}$ are endpoint relative permeability at high and low capillary numbers, respectively. Now, using a simple function such as Corey-type relative permeability equation, we can calculate relative permeability.

$$k_{rl} = k_{rl}^0 (\overline{S_l})^{n_l} \tag{2.14}$$

where $\overline{S_l} = \frac{S_l - S_{lr}}{\sum_{l=1}^3 S_{lr}}$ is the normalized saturation and n_l is the Corey exponent for l^{th} phase.

2.9 Conductivity measurement

As stated in section 1.5, to account for the right degree of mixing in the heterogeneous reservoir, dispersivity must be considered when the core is modeled as a homogeneous core. In literature, electrical conductivity (EC) is found to have a linear relationship with the tracer concentrations (TDS) [61, 62], which is given as TDS = A x EC. This can be confirmed from the figure 2.9(a), where the conductivity can be seen to increase linearly with TDS in the range of 9400-20000 ppm salinity. To obtain the right value of dispersivity, we perform a separate experiment where the conductivity of the effluents in the extracted oil is measured during the initial water flooding (IWF) and high TDS slug (HTDS) since the oil production is steady during these injection stages. Figure 2.9 shows that using the same value of dispersivity (0.01ft) for all the three phases i.e. aqueous, oleic, and microemulsion, gives

an excellent match with the experimental results. This is also in agreement with the literature on Berea sandstones [63].

(a)

(b)
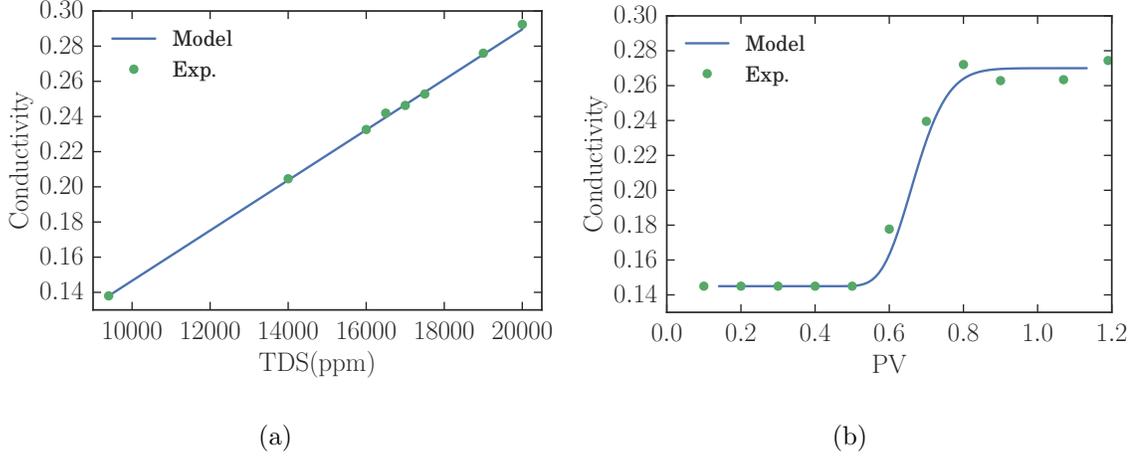
Fig. 2.9. a) Linear relationship between conductivity and TDS solution. b) Validation of conductivity calculated from the resulting linear relationship for a coreflood experiment with dispersivity value 0.01 ft.

2.10 Adsorption model for surfactant and polymer

The effect of adsorption is assumed to be irreversible with concentration and reversible with salinity. Figure 2.10 shows the mechanism of polymer or surfactant adsorption on a rock surface where the residual adsorption increases with increase in concentration and for a certain value of concentration reaches a maximum constant value. In case of SP flooding, surfactant and polymer molecules are competitively adsorbed [64]. We use a Langmuir type isotherm to model this adsorption curve for both the surfactant and polymer:

$$\hat{C}_m^* = \frac{a_m(\tilde{C}_m - \hat{C}_m^*)}{1 + b_m(\tilde{C}_m - \hat{C}_m^*)}, \quad m = 1, 2 \tag{2.15}$$

where

$$a_m = (a_{m1} + a_{m2}C_{SE})(\frac{k_{ref}}{k})^{1/2} \tag{2.16}$$
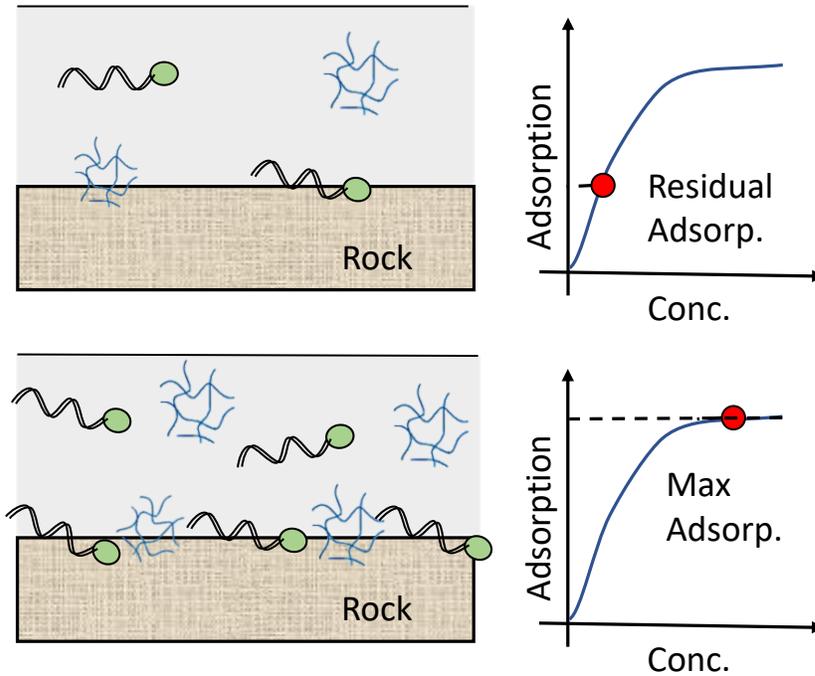
Fig. 2.10. Adsorption of surfactant and polymer in reservoir. Adsorption increases with increase in concentration which then reaches a constant maximum value. The resulting curve is modeled using Langmuir isotherm equation.

where $m = 1$ is surfactant and $m = 2$ is polymer, a_{m1}, a_{m2} and b_m are fitting parameters, C_{SE} is the effective salinity, $\hat{C}_m^*$ is volume of adsorbed surfactant or polymer/volume of water, and $\tilde{C}_m$ is adsorbed surfactant or polymer concentration/pore volume.

2.11   Final model calibration

In this section, we present the final model calibration values obtained for different models using extensive laboratory scale experiments (refer to table 2.4). Interested readers may read [27] which explains in detail the significance of all parameters shown in table 2.4.

Table 2.4.

Final SP flood model parameter set obtained after initial model calibration stage. Adapted from [65].

UTCHEM parameter(s)	Description	Value(s)
Polymer models		
AP1, AP2, AP3	Zero shear viscosity multipliers	15, 0, 155
BETAP	Effective divalent salinity parameter	1
CSE1	Salinity above which viscosity is affected	0.0001
SSLOPE	Log-log slope of viscosity-effective salinity	-0.76
GAMMAC,POWN	Polymer shear rate dependency parameters	40, 1.66
A0, A1, A2	Constants in Eq. (2.1)	$-588, 7683, -12$
Surfactant models		
C2PLC, C2PRC	Oil concentrations at the plait point of type II ($-$), and type II ($+$)	0, 1
EPSME	Critical micelle concentration	0.0001
CSEL,CSEU	Salinity window for type III	0.255, 0.305
HBNC70, 71, 72	Heights of binodal curve at zero, optimal, and twice optimal salinities	0.027, 0.0155, 0.02
chuh, ahuh	Constants in Chun Huh IFT model	0.3, 10
Xiftw	Log of oil-water interfacial tension	1.32
ALPHAVi; i: 15	ME viscosity parameters	2, 2, 0, 0.9, 0.7
Chemical-rock interactions		
t11, t22, t33	CDC parameters for the aqueous, oleic, and ME phases	1500, 1100, 346
e1c, e2c, e3c	Corey exponents at high Ca numbers for the aqueous, oleic, and ME phases	2.0, 2.0, 2.0
AD31, AD32, B3D	Surfactant adsorption parameters	0.31, 0, 100
AD41, AD42, B4D	Polymer adsorption parameters	8.4, 0, 100
EPHI4,BRK,CRK	Polymer permeability reduction parameters	0.85, 100, 0.068

2.12   BCF experiments

As stated previously, as opposed to similar studies where only the overall oil recovery is history matched [66], our aim is to history match all the BCF experiments for two quantities of interests. Among many factors determining the fate of SP flooding, chemical composition and specifically salinity plays a critical role as presence of surfactant creates a complex phase behavior of oil/water/surfactant which is greatly affected by the slug design and salinity [67]. Thus, the BCF experiments are all performed at different salinities (TDS) and slug design (reported in pore volume (PV)) to study these effects on oil recovery. Injection for all BCF experiments follows

Table 2.5.
Injection sequence and data of BCF experiments. Concentrations are in ppm.

Slug	Property	BCF 1	BCF 2	BCF 3
IWF	PV	1.45	1.60	1.60
	TDS	9400	9400	9400
HTDS	PV	—	1.0	0.25
	TDS	—	17500	17500
SP	PV	0.50	0.25	0.25
	TDS	9400	17500	17500
	Surf. conc.	6780	6780	6780
	Poly. conc.	3300	2500	2500
P	PV	1.01	0.50	0.51
	TDS	9400	9400	9400
	Poly. conc.	3300	2500	2500
EWF	PV	1.27	1.59	1.49
	TDS	9400	9400	9400

the same sequence including initial water flood (IWF) $\rightarrow$ high TDS (salt) preflush

(HTDS) → surfactant-polymer flood (SP) → polymer flood (P) → extended water flood (EWF). The details of these BCF experiments is given in Table 2.5 and can also be found in [65]. Our aim is to determine a set of model input parameters that can perfectly predict oil-cut profile, pressure drop profile (i.e. pressure difference between injection and production well ($|P_{inj} - P_{prod}|$) and cumulative oil production profile for all BCF experiments without any ad-hoc tuning.

In the next section, we quantify the uncertainty in BCF experiments due to uncertainty in the model parameters. The intrinsic variability in the physical properties governing the mechanism of oil recovery is an inherent characteristic of any geophysical system. Successful design of any EOR strategy depends heavily on how well this uncertainty due to parameter variability is captured. Therefore, we consider $\pm 20\%$ uncertainty in every physical parameter as a prior. Then we perform 1000 simulations for each BCF experiment and quantify the uncertainty in the output response of our model due to uncertainty in these model parameters (refer to figures 2.11-2.13). The results show a great uncertainty in the output response for all the BCF experiments, therefore signifying a need for automatic history matching algorithm. In the next chapter, we present a sequential algorithm that uses sensitivity analysis, proxy modeling and inverse optimization to successfully history match all BCF experiments without any ad-hoc tuning of these model parameters.

(a) Pressure drop   (b) Cumulative oil recovery   (c) Oilcut
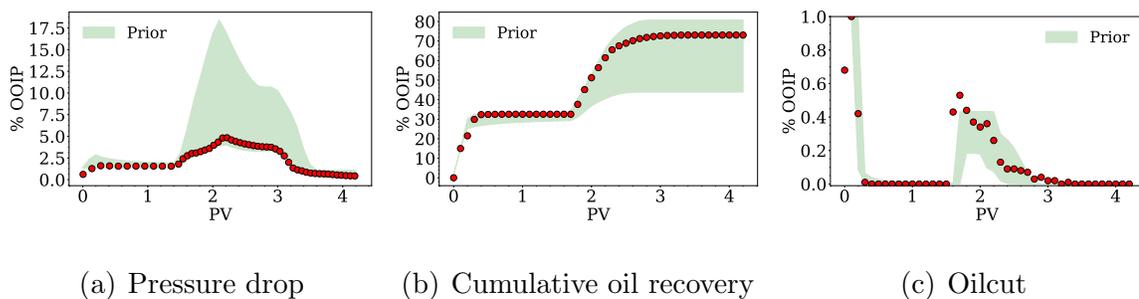
Fig. 2.11. The resulting prior space for BCF1 obtained using initial model parameter set provided in table 2.4 for a) pressure drop profile b) cumulative recovery curve c) oilcut curve.

(a) Pressure drop   (b) Cumulative oil recovery   (c) Oilcut
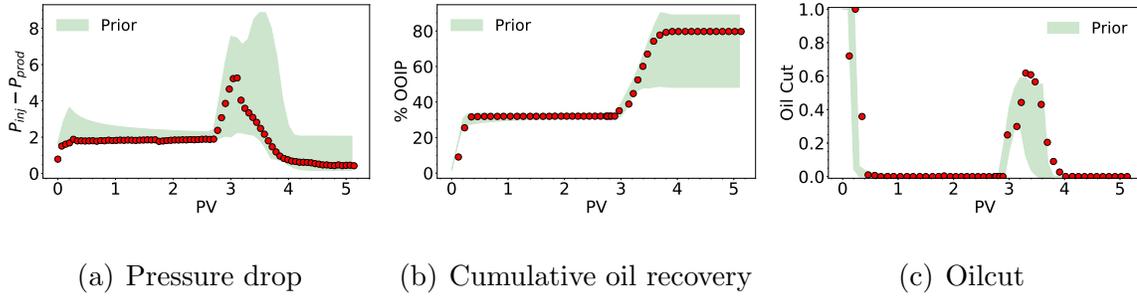
Fig. 2.12. The resulting prior space for BCF2 obtained using initial model parameter set provided in table 2.4 for a) pressure drop profile b) cumulative recovery curve c) oilcut curve.

(a) Pressure drop   (b) Cumulative oil recovery   (c) Oilcut
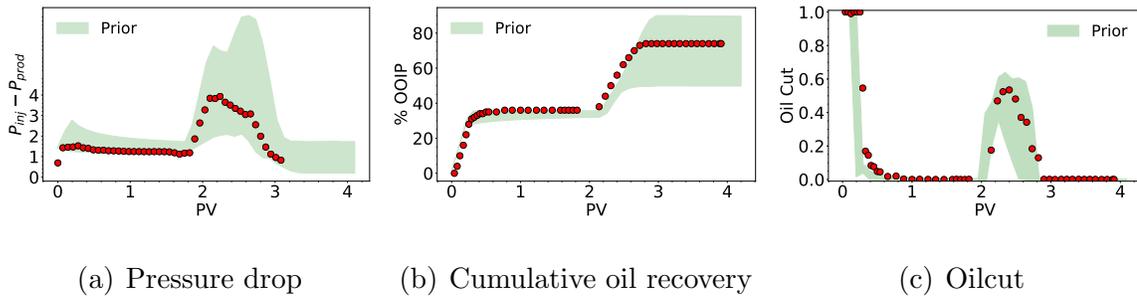
Fig. 2.13. The resulting prior space for BCF3 obtained using initial model parameter set provided in table 2.4 for a) pressure drop profile b) cumulative recovery curve c) oilcut curve.

2.13   Conclusion

In this chapter, we presented a comprehensive review on the initial model calibration stage of SP flood process. We first constructed a mechanistic model of SP flood process by dividing the SP flooding into three sub-processes. Then using extensive laboratory experiments such as polymer rheology, surfactant phase behavior, polymer permeability reduction, and capillary desaturation, we performed initial model calibration. We improved the polymer shear-thinning viscosity model in UTCHEM

to account for variation of half shear rate ($\dot{\gamma}_{1/2}$) with concentration and salinity and showed that it results in a better agreement with the experimental results.

We then considered $\pm 20\%$ uncertainty in the calibrated parameters to quantify uncertainty in BCF experiments. The results showed a great uncertainty in the output response of simulation model for all our quantities of interests such as overall oil recovery profile and pressure difference curve. In the next chapter, we aim at quantifying this uncertainty in a systematic way and present a sequential assisted history matching algorithm that calibrates all BCF experiments without any ad-hoc tuning of model parameters.

3. HISTORY MATCHING USING POLYNOMIAL CHAOS-BASED PROXY

A version of this chapter has been published as:

Pratik Naik, Soroush Aramideh, and Arezoo M. Ardekani. "History matching of surfactant-polymer flooding using polynomial chaos expansion." *Journal of Petroleum Science and Engineering* (2018).

3.1 Abstract

In this chapter, we proposes a robust framework for history matching which employs a sequential execution of sensitivity analysis, proxy modeling and inverse optimization to determine the optimized parameter space of model parameters. The mechanistic surfactant-polymer (SP) flood model developed in Chapter 2 is considered for history matching with an ultimate goal of accurately calibrating models that describe physical subprocesses of surfactant flooding, polymer flooding and displacement process. The employed model calibration algorithm starts with Sobol sensitivity analysis which reduces the large uncertain space of model parameters to determine the most important stochastic variables. The resulting low-dimensional parameter space is then represented via appropriate orthonormal basis of polynomial chaos expansion (PCE-proxy). An inverse optimization problem is then posed that minimizes the miss-fit between PCE-proxy response and experimental observations by employing a Genetic Algorithm. Finally, the epistemic uncertainty in PCE-proxy is quantified by combining it with a Gaussian regression process called Kriging.

We use this framework to calibrate the SP flood model by history matching a single coreflood experiment for quantities of interest such as pressure drop profile and cumulative oil recovery curve. We then show that the calibrated model is success-

fully able to predict all our quantities of interest for two other coreflood experiments without any ad-hoc tuning of parameters. The proposed proxy-accelerated inverse optimization framework shows significant promise for model calibration or to improve the quality of history matched results.

3.2 Introduction

History matching is an inverse problem where reservoir properties and/or flow physics models are calibrated by matching the predictions of mathematical models to observed reservoir data. History matching problems, however, falls under the category of ill-posed problems, meaning there exists many solutions that could predict the observed reservoir history data equally well [?]. In the last two decades, advancement in computational capabilities [68] has led to a drastic shift from manual history matching methods to automatic or assisted history matching methods. In manual history matching methods, the reservoir parameters are tuned manually to predict reservoir history [69, 70], while assisted history matching methods make use of advance techniques such as stochastic optimization algorithms [71], probabilistic approach [72] or proxy modeling [73] to determine the best model parameter set for accurate prediction of the observed data. Assisted history matching methods have given a new outlook to history matching process, where instead of obtaining a single correct set of model parameters, the focus is on quantifying uncertainties in stochastic model parameters by simultaneously matching multiple observed data [74, 75]. A comprehensive review on different history matching methods is provided in [?, 76–79].

Stochastic optimization algorithms such as Genetic Algorithm [71], Particle Swarm Optimization [80] or Evolutionary Strategies [81] have been well-studied and successfully implemented for reservoir history matching. One of the main drawback of such stochastic methods is a slow convergence rate which requires thousands of evaluation of forward simulation model resulting in prohibitive computational cost. An alternative is to use a Bayesian inference which is a probabilistic approach for history

matching [82]. In Chapter 4, we discuss the implementation of Bayesian approach for history matching. Bayesian uses information on priors of model parameters to maximize the likelihood of the observed data and quantifies uncertainties in these model parameters with a posterior distribution. However, for a high dimensional problem such as surfactant polymer flooding, the posterior distribution is generally multi-mode non-Gaussian, requiring efficient sampling strategies [83]. One of the most commonly used strategy to sample from posterior distribution is Markov Chain Monte Carlo (MCMC). MCMC, however, surfers from high rejection rates and like any other stochastic process requires forward reservoir model evaluations for every proposed sample [84]. Even though various improvements have been proposed recently to the traditional MCMC sampling methods such as hybrid MCMC [85], population MCMC [86], or ensemble Kalman Filter [87], evaluating flow simulation models even for few thousands of proposed MCMC samples are impractical for a full-scale reservoir model. This has warranted the use of surrogate models commonly referred to as 'proxy', which approximates the non-linear response of forward simulation model for selected input parameters and facilitates a significant reduction in computational time.

There exists a number of successful implementation of proxy modeling for reservoir history matching in the literature such as surface response modeling [88, 89], PCE surrogate modeling [90] or artificial neural networks [91, 92]. A detailed comparison of different proxy-modeling techniques in reservoir history matching applications can be reviewed in [93, 94]. The choice of proxy model depends on the application and understanding of the problem at hand. In this work, we have chosen Polynomial Chaos Expansion (PCE) as our proxy model because of these reasons: i) it can be efficiently constructed in a non-intrusive way by treating the simulation model as a 'black-box'without any modification to the governing equations ii) there is no need of efficient sampling algorithm to model the non-linear behavior of output response (as in Bayesian) and most importantly iii) theoretically guarantees convergence of output random variables as the degree of PCE is increased.

PCE has been extensively applied as a proxy in the area of stability and control [95,96], solid mechanics [97,98], electronic circuits [99,100] or computational fluid dynamics [101]. The origin of PCE theory dates to 1938 when Wiener first introduced homogeneous chaos in his paper [102]. In this work, Wiener first extended Birkhoff ergodic theorem from one-dimensional variable to multidimensional truncated Hermite functions and then showed that Brownian motion theory can be used to model stochastic processes by describing random variables as Gaussian distributions. In 1946, Cameron and Martin [103] developed a theory to express any non-linear functional as a series of Hermite-Fourier functional by introducing set of orthonormal functionals on a Wiener process and showed that this resulting series converges in a mean square (L^2) way. Meechan in 1964 [104,105], expanded Wiener-Hermite expansion theory on a one-dimensional Burgers turbulence model. Meechan argued that the field variables of turbulence can be approximated as a series with first term as Gaussian distribution and remaining terms as correction to Gaussian distribution. However, Orszag in 1967 [106] showed that (2nd order) Wiener-Hermite truncation of Burgers model suffers from closure problem mainly because of non-uniform convergence of Wiener-Hermite expansion. They proved that inviscid equipartition solutions fail in closure and hence, higher order terms are necessary to represent shocks in a high Reynolds number turbulent flow which implies a slow convergence rate. Crow [107] in 1970, resolved this issue by showing that time varying Lagrangian basis could be adopted to speed up the convergence rate. However, limited development on the applicability of Wiener-Hermite expansion to spectral stochastic methods resulted in less popularity of this method in 1970s and 1980s. One noticeable work during this period was of Ogura [108] who demonstrated the use of Charlie polynomial chaos for orthogonal functional of Poisson process. Then, Ghanem and Spanos [97] in 1991, presented a formulation using truncated Hermite polynomials that could accommodate random functions in the Galerkin framework of deterministic finite element methods. They in a way revolutionized the use of polynomial chaos by introducing it to engineering applications of solid mechanics for the first time. Xiu [109] extended the work

of Ghanen and Spanos, and Ogura to propose Wiener-Askey polynomial chaos framework for stochastic processes and showed that the use of optimal trial basis leads to exponential convergence of error. Their work was pivotal in extension of polynomial chaos theory to arbitrary random distribution of stochastic input variables [110].

In this chapter, we are particularly interested in the application of PCE technique for proxy modeling and history matching in reservoir simulations. Sharma's work [73] was probably one of the first that implemented PCE for real time optimization and uncertainty propagation in reservoir simulations. They proposed a closed loop approach where Karhunen-Loeve (KL) expansion and Bayesian inversion were used for real-time history matching. Since then, many studies have been published where researchers have exploited the use of PCE as a proxy model for uncertainty quantification [111], optimization [94] or global sensitivity analysis [112, 113]. Also, many improvements have been proposed to the traditional PCE technique. Camacho [111] first proposed Smolyak quadrature method to calculate PCE coefficients and then compared the accuracy of this new method with regression method, while Sarma [90] used a non-intrusive spectral projection method to calculate PCE coefficients. So far, the focus of all these studies has been to calibrate reservoir physical parameters such as permeability, porosity, residual water saturation, etc. To the best of our knowledge, there has not been an attempt to calibrate flow physics models of surfactant-polymer flooding process in a systematic way using PCE-proxy. Some of the major challenges in calibrating SP flooding process using a proxy-modeling approach are i) lack of laboratory core flood experimental data to calibrate physical models ii) high dimensionality of the input space because of many fitting parameters associated with different models of each coreflood process iii) problems associated with the accuracy, efficiency and uncertainty in proxy model.

In this work, we make an attempt to address all these problems. Here, we present a framework for history matching and model calibration of surfactant-polymer (SP) coreflood experiments where cumulative oil recovery curve and pressure drop profile are history matched for three different coreflood experiments. The methodology pre-

sented involves sequential execution of sensitivity analysis, surrogate modeling and global optimization. At first, a mechanistic model of SP flood is developed by dividing the entire process into three sub-processes separately: i) surfactant flooding ii) polymer flooding and iii) displacement process. Then, different models describing each of these processes are calibrated using extensive laboratory experiments including polymer rheology, surfactant phase behavior, polymer permeability reduction, and capillary desaturation experiments. Refer to Chapter 3, to learn more about the initial model calibration stage.

Initial model calibration results in a large set of input parameters (upto 19), if used directly for proxy construction, reduces its efficiency. Hence, Sobol based global sensitivity analysis (SA) is performed on these parameters to determine the most important ones with respect to overall oil recovery factor (reported as original oil in place or %OOIP) and maximum pressure drop ($max(\Delta p)$) which are the main quantities of interest in real applications. A PCE-proxy is then constructed to map these stochastic input parameters obtained from SA to stochastic output predictions. Predictive capabilities of PCE-proxy are tested by performing rigorous validations (with $N_s = 50, 200, 500, 1000$ evaluations of simulation model). The robustness of PCE-proxy is also evaluated by comparing it with benchmark Monte Carlo simulations (10^4 samples) performed with Latin Hyper Sampling (LHS). The accuracy of PCE-proxy is quantified in terms of relative generalization error. Finally, an inverse problem is formulated to find the optimized model parameters by minimizing the misfit between observed and predicted data with a Genetic Optimization Algorithm (GA). At last, the PCE model is extended to incorporate a Kriging model to capture the uncertainty in the proxy-model itself. In other words, probable errors associated with the predicted values of proxy are measured using PCE-Kriging (PCE-K).

This chapter is structured as follows. In section 3.3, we briefly describe SP coreflood experimental setup used in this work and present the methodology that has been adopted to calibrate its models. Section 3.4 presents sensitivity analysis framework that facilitates an accurate and robust PCE-proxy construction. Section 3.5 provides

the theory and details of PCE-proxy. This section also provides error estimates in PCE-proxy posterior and its validation with LHS simulations. In section 3.6, we pose an inverse optimization problem to minimize the miss-fit between observed data and PCE-proxy output response. Finally, section 3.8 discusses implementation of Kriging process to quantify epistemic uncertainty in PCE-proxy itself.

3.3   Framework for Model Calibration

We perform experiments on Berea corefloods (BCF) with an objective to study the importance of chemical composition (salinity, surfactant and polymer concentration) and slug sizes on overall oil recovery. The details of these BCF experiments are given in Table 2.5 and can also be found in [65]. In this chapter, we aim to determine a set of model input parameters that can perfectly predict oil-cut profile, pressure drop profile (i.e. pressure difference between injection and production well ($|P_{inj} - P_{prod}|$) and cumulative oil production profile for all BCF experiments without any ad-hoc tuning.

We follow a sequential workflow for history matching process as depicted in figure 3.1. The methodology presented can be considered to be separated into three stages: sensitivity analysis, proxy construction and an optimization, which are sequentially executed. Their sequential execution allows us to use strengths of each of these stages to overcome the drawbacks of the following stages.

The first stage, sensitivity analysis, is aimed at reducing the dimensions of stochastic model input vector (X) to make the proxy-modeling efficient and robust. Algorithm 1 (refer to Appendix A) explains the steps involved in sensitivity analysis (SA), where the Sobol method is used to find the most important input parameters (ξ) ($X \to \xi$). This reduced parameter set is then used to construct an efficient proxy model using PCE ($P(\xi)$) as explained by Algorithm 2 (refer to Appendix A) ($\xi \to P(\xi)$). Recently, Rohmer et al. [114] employed a similar approach using variance based global sensitivity to reduce the dimensionality of their response
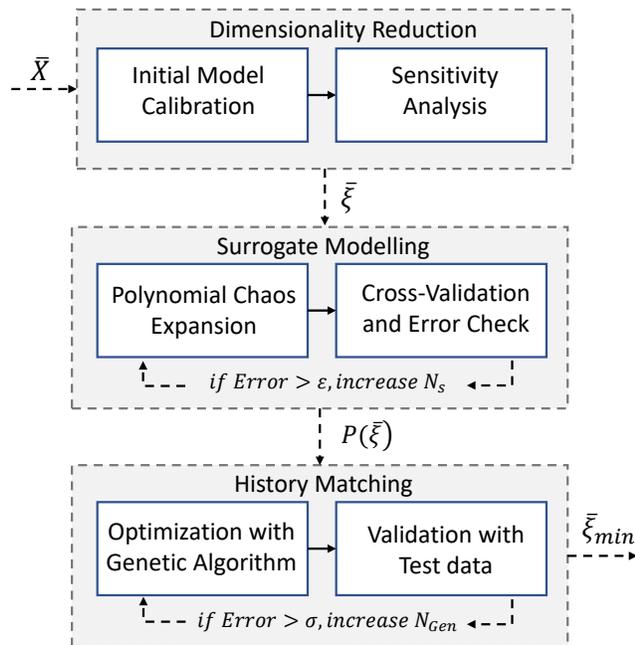
Fig. 3.1. Proposed workflow for history matching using PCE-proxy. The workflow is divided into three stages which are executed sequentially as a) dimensionality reduction b) surrogate modeling c) inverse optimization.

surface-based proxy model to study CO_2 storage in reservoirs while Bazargan [68] used Karhunen-Loeve expansion method as dimensionality reduction component to construct a regression based PCE-proxy for quantifying uncertainty in their reservoir permeability field. Also, Alejandra Camacho et al. [111] showed that the PCE-proxy model constructed using optimal design based approach performs better as compared to computationally intensive Monte Carlo simulations.

Next, we use the constructed proxy model with an optimization algorithm to obtain an optimized parameter set (ξ_{min}). In this optimization stage, we start with a proposed set of initial solutions based on support-experiments and our physical understanding of the problem, and use optimization to identify the optimized model

parameters using BCF1 ($P(\xi) \to \xi_{min}$). This involves defining a multi-objective inverse model calibration problem where the objective is to minimize the miss-fit between observed values and PCE response for pressure drop profile and cumulative oil recovery curve. Recently, Ampomah et al. [115] adopted a similar multi-objective optimization approach using polynomial surface response proxy to optimize both the oil production and CO_2 storage in a mature oil reservoir. Such proxy-based optimization methods have been extensively studied [116–118].

We then test the predictive capability of our approach by showing that the identified input parameters work equally well for BCF2 and BCF3. This is explained by Algorithm 3 (refer to Appendix A) where a Genetic Algorithm (GA) is used to find a set of perfect history matched models [119]. Finally, we quantify the uncertainty in our proxy model by performing a Kriging process [120].

3.4   Parameterization and Sensitivity Analysis

Parametrization is the process of representing any high-dimensional model in terms of physically interpretable quantities that greatly affects the output response of the model [68]. We parameterize the problem at hand by dividing the SP flooding process into three sub-processes (surfactant flooding, polymer flooding and displacement process) and then modeling each of these sub-processes with different physical models (refer to Chapter 2). Table 3.1 shows all the different models used in this work corresponding to each core-flood process and their fitting parameters (upto 19 in total). For such high dimensional problems with non-linear response, it becomes necessary to reduce the dimensions of uncertain input parameters for efficient proxy model construction. This can be achieved through SA, where the emphasis is on finding the most important input parameters with respect to output responses. Here, we employ a Sobol based global sensitivity analysis [121, 122].

In Sobol sensitivity analysis, a functional analysis of variance (V) decomposition method (ANOVA) is used to rank the parameters based on the effects of input vari-

ability on the output variability of a model response. Suppose, we have a model function f which maps:

$$f : X \to Y, \tag{3.1}$$

where $X = (X_1, X_2...X_n)^T$ is the model input vector and $Y = f(X)$ is the model output response. Sobol indices (rank) can then be calculated as:

$$S_i = \frac{V_{Xi}[\mathbb{E}_{X\sim i}(Y|X_i)]}{V(Y)}, \tag{3.2}$$

where the terms of functional decomposition are expressed in terms of conditional expected values ($\mathbb{E}$). The sum of all Sobol indices, called total Sobol indices (S_i^T), can then be estimated as follows:

$$S_i^T = \sum_{k\neq i} S_k = \frac{\mathbb{E}_{X\sim i}[V_{Xi}(Y|X_{\sim i})]}{V(Y)}. \tag{3.3}$$

In this study, we separately analyze each model by quantifying the model output response with respect to overall oil recovery factor and maximum pressure drop ($max(\Delta p)$). Accurate prediction of $max(\Delta p)$ is of practical importance as unwarranted pressure changes may lead to stress changes in the reservoir, potentially affecting the fracture permeability or often leaking fluids through the micro-fractures [123,124]. Thus, $max(\Delta p)$ is used as a second objective function in sensitivity analysis in order to build a more comprehensive surrogate model and accurately predict pressure drop. We perform 1000 simulations on each model to identify the most sensitive parameters within each model. It has been tested that 1000 simulations are enough to produce converged results. Figure 3.2 shows the results of sensitivity analysis on the final overall oil recovery factor and $max(\Delta p)$ with respect to total Sobol indices given by Eq. (3.3).

A more detailed physical interpretation of all the variables characterizing the models in Table 3.1 can be found in [27]. It can be seen from figure 3.2 that for the surfactant phase behavior model, C_{33max0} and C_{33max1} are the most sensitive variables. For the adsorption model, we use a Langmuir isotherms to describe the adsorption level, where we assumed that the salinity has not effect on adsorption (i.e.

Table 3.1.
Models used for Sensitivity Analysis and their model fitting parameters.

Sub-process	Model Name	Model Parameters
Surfactant Flooding	Micro-emulsion Viscosity	$\alpha_{V1}, \alpha_{V2}, \alpha_{V3}, \alpha_{V4}, \alpha_{V5}$
	Phase Behavior	$C_{33max0}, C_{33max1}, C_{33max2}$
	Adsorption	ad_{31}, ad_{32}, b_{3d}
Polymer Flooding	Viscosity and Shear Rate	$\dot{\gamma}_{1/2}$
	Permeability Reduction	c_{rk}, b_{rk}
Displacement Process	Relative Perm. (Corey Model)	n_{high}, n_{low}
	Dispersivity	$\alpha_{L1}, \alpha_{L2}, \alpha_{L3}$
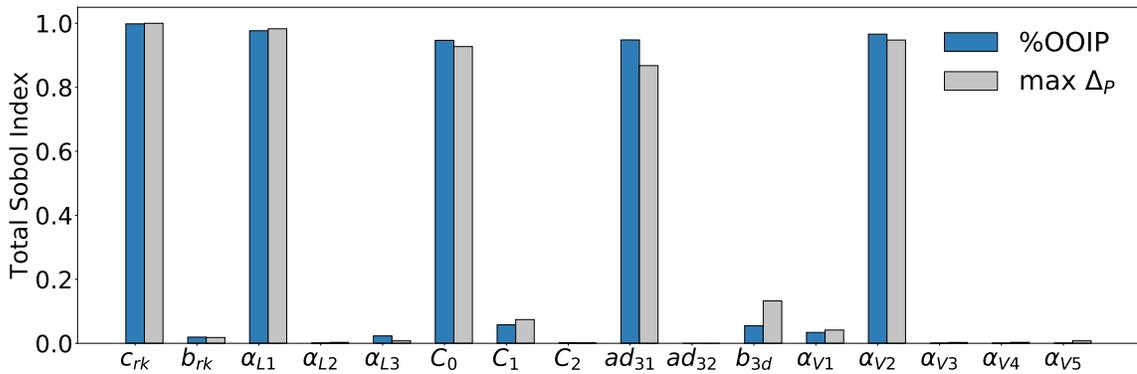
Fig. 3.2. Sensitivity analysis on different models of SP flooding using the Sobol technique. (Here, C_0, C_1, C_2 represents $C_{33max0}, C_{33max1}, C_{33max2}$ respectively).

ad_{32} is kept 0). The fitting parameter b_{3d} controls the curvature of Langmuir isotherm curve while the ratio (ad_{31}/b_{3d}) controls the peak of adsorbed surfactant. Here, ad_{31} shows the most sensitivity. For the displacement process, we include longitudinal dispersivity in our 1D simulation model to capture the effects of heterogeneity. In this dispersion model, the uncertainty in α_{L1} shows the most sensitivity which models the dispersivity of aqueous phase. The selected fitting parameters from SA that are

used to construct proxy-model are summarized in Table 3.2. We have performed all the experiments at the room temperature which restricted the calibration of our models to capture the effects of temperature variation. However, interested readers may refer to [125,126] which study the effects of temperature variation on surfactant and polymer flooding.

Table 3.2.
Range of the input parameters for polynomial chaos proxy model and their sources.

Model Name	Sensitive Par.	Min	Max	Source
Micro-emulsion Viscosity	α_{V2}	0	3	[127] [128] [4] [129]
Phase Behavior	C_{33max0}	0.02	0.05	Experiment [60]
	C_{33max1}	0.01	0.02	Experiment
Adsorption	ad_{31}	0	2	Experiment
Viscosity and Shear Rate	$\gamma_{1/2}^{\cdot}$	55	100	Experiment
Permeability Reduction	c_{rk}	0.05	0.1	Experiment
Relative Perm. (Corey Model)	n_{high}, n_{low}	1.5	3	[130]
Dispersivity	α_{L1}	0.008	0.012	Experiment [131]

Before we begin our proxy-model design, a prior range needs to be defined on all the model input parameters. Parametrization should honor the model-constraints as well as any geological constrains and hence only a reasonable bound should be used on these parameters. The prior ranges of these parameters are determined either by conducting support-experiments [65] and by considering an uncertainty of $\pm 20\%$ on these model fitting parameters which is a reasonable assumption in reservoir problems [132]. The priors on those parameters which are not possible to infer from these experiments are determined after doing an extensive literature review (refer to Table 3.2). It is worth noting that these same bounds on model parameters will be used for the optimization problem in section 3.6.

3.5 Polynomial Chaos Expansion

This section explains the construction of non-intrusive PCE-proxy model. Here, we adopt the framework from UQLab [133]. Consider a reduced order vector of $\mathbf{n}$ dimensions of uncertain input variables (given in Table 3.2), denoted by $\xi = (\xi_1, \xi_2...\xi_n)$. Let the output response of our finite variance simulation model be denoted by $\mathbf{Y}$. Our quantities of interest (QoIs) are pressure drop profile and cumulative oil recovery curve. We seek to approximate the output response $\mathbf{Y}$ of our simulation model with a known polynomial function $P(\xi)$. This can be approximated using the weighted sum of orthonormal polynomials as [97, 109]:

$$Y \approx P(\xi) = \sum_{\alpha \in A} y_\alpha \Psi_\alpha(\xi) + \epsilon, \tag{3.4}$$

where $\Psi_\alpha(\xi)$ are a multivariate orthonormal basis functions with respect to joint probability density (PDF) $f(\xi)$ of input variables. These are denoted by multi-index $\alpha \in A$ where $(\alpha_1, \alpha_2...\alpha_p)$ are nodes on which the sum is approximated and A is the truncation set of degree p. y_α are the coefficients that we seek to determine by evaluating the simulation model at these nodal points. The truncation error of PCE is given by ϵ. Therefore, if the polynomial function $P(\xi)$ is accurately constructed with minimal truncation error ϵ, it can be used as a proxy model to approximates the actual output response of simulation model.

The multivariate orthonormal basis functions $\Psi_\alpha(\xi)$ are obtained from tensor product of univariate orthonormal polynomial ϕ_{α_i}. We consider a *uniform* probability distribution $f(\xi) \sim \mathcal{U}(\prod_{j=1}^n [a_j, b_j])$ for all the input variables for which the set of orthonormal basis functions are *Legendre* polynomials. The lower bound a_j and upper bound b_j of the distribution for each parameter is provided in table 3.2. The polynomial function given by Eq. (3.4) is generally approximated upto p degree by using different truncation schemes ($A^{n,p} = [\alpha \in \mathbb{N}^n : |\alpha| \le p]$). We disregard the higher order interacting terms in the approximation using q-norm or hyperbolic scheme, where $0 < q < 1$ is a parameter that determines the hyperbolic truncation surface and is given by $|\alpha| = (\sum_{i=1}^n \alpha_i^q)^{1/q} \le p$ [134].

The coefficients of the polynomial function given by Eq. (3.4) can be determined with either an intrusive approach (such as Galerkin), where the equations governing the physical model are modified, or with a more direct alternative called non-intrusive approach, which requires evaluation of simulation model at experimental design points (projection method [135], stochastic collocation method [136] or least square regression method [137]). Here, we use Least Square Minimization approach to determine these coefficients. An advantage of this method is that any arbitrary set of points that represents the true probability distribution ($f(\xi)$) of input variables can be used to determine the coefficients.

3.5.1 Least Square Minimization to determine coefficients

In least square minimization, we seek to minimize the mean square truncation error in Eq. (3.4) as [138]:

$$y_\alpha = \underset{y_{\alpha \in R^{P_{basis}}}}{arg\ min} \frac{1}{N_s} \sum_{i=1}^{N_s} \left(P(\xi^{(i)}) - \sum_{\alpha \in A} y_\alpha \Psi_\alpha(\xi^{(i)}) \right)^2, \tag{3.5}$$

where N_s are the total number of experimental design points generated from probability distribution of input variables and $\Xi = (\xi^{(1)}, \xi^{(2)}...\xi^{N_s})$ is the input vector. We generate this sample points by using Latin Hypercube Sampling (LHS). Numerous studies show that LHS has improved convergence rate over Monte Carlo Sampling (MC) method especially in multidimensional cases [139–141]. The computational model is evaluated on each of these points. An Ordinary Least Square (OLS) method, which minimizes the squared residual sum, can be used to solve Eq. (3.5), which reads [142]:

$$y_\alpha = (A^T A)^{-1} A^T \mathbb{Y}. \tag{3.6}$$

where $\mathbb{Y} = (y^{(1)}, y^{(2)}...y^{N_s})$ is output model response evaluated at N_s size of input vector and A is called experimental design matrix that contains orthonormal bases in all N_s and can be written as $A_{ij} = \psi_j \xi^{(i)}$, $i = 1...N_s, j = 1...p - 1$

Since Eq. (3.5) is a least square minimization problem, an increase in the number of experimental design points (N_s) results in minimization of truncation error, which inturn translates to a more robust and accurate PCE-proxy. However, evaluating the simulation model at large number of experimental design points is computationally expensive and hence it is important to estimate errors in PCE posterior predictions to determine the optimum number of model evaluations.

3.5.2 PCE posterior error estimates

The error in posterior of PCE can be estimated quite accurately by the means of least square minimization error also known as relative *generalization* ϵ_{gen}. One of its variant known as *normalized empirical error* ϵ_{emp} quantifies the predictive ability of PCE to reproduce the output response on experimental design points (Ξ) as [143]:

$$\epsilon_{emp} = \frac{\sum_{i=1}^{N_s} \left(P(\xi^{(i)}) - P^{PC}(\xi^{(i)}) \right)^2}{Var(\mathbb{Y})}. \tag{3.7}$$

where $Var(\mathbb{Y}) = \sum_{i=1}^{N_s} \left(P(\xi^{(i)}) - \frac{1}{N_s} \sum_{i=1}^{N_s} P(\xi^{(i)}) \right)^2$ is variance of model output response of $\mathbb{Y}$. ϵ_{emp}, however, often leads to over-estimation of true error as it does not evaluate or cross validate the performance of PCE on points other than experimental design points [144]. One way to overcome this issue is dividing the experimental design set into two random sets: building the PCE with one set and cross-validating the PCE performance on the other set. One such measure of error is *leave-one-out error* (ϵ_{LOO}), where PCE is constructed on all but one point ($P^{PC\sim i}(\xi)$), while relative error in PCE is evaluated on that excluded point (i) [145]. This process is then repeated for all points in the experimental design set N_s [146].

$$\epsilon_{LOO} = \frac{\sum_{i=1}^{N_s} \left(P(\xi^{(i)}) - P^{PC\sim i}(\xi^{(i)}) \right)^2}{Var(\mathbb{Y})}. \tag{3.8}$$

It is important to note that, the output response of our simulation model ($\mathbb{Y} = (y^{(1)}, y^{(2)}...y^{N_s})$) is time dependent. So, $y^{(i)}$ is a vector response discretized at k steps

and is given by $y^{(i)} = y_j^{(i)} = (y_1^{(i)}, y_2^{(i)}...y_k^{(i)})$ where $j = 1...k$. So, we define ϵ_{LOO}^{mean} (or ϵ_{emp}^{mean}) as average of LOO error (or empirical error) at all k steps.

$$\epsilon_{LOO}^{mean} = \frac{\sum_{i=1}^{k} \epsilon_{LOO}^i}{k}. \tag{3.9}$$

$$\epsilon_{emp}^{mean} = \frac{\sum_{i=1}^{k} \epsilon_{emp}^i}{k}. \tag{3.10}$$

ϵ_{LOO}^{mean} or ϵ_{emp}^{mean} error captures the overall error in the PCE output response $(y^{(i)})$. In this work, we discretized the output response of our simulation model such that $y_k^{(i)} - y_{k-1}^{(i)} = 0.1$ PV.

3.5.3 Experimental design size selection and error estimates

We start by investigating the effect of increasing experimental design sample size (N_s) on the accuracy of PCE prediction. We use $\Delta\epsilon_{LOO}^{mean} = 5 \times 10^{-4}$ as stopping criteria in Algorithm 2 (line 2). The degree of polynomial is given a range of $1 \le p \le 5$ with hyperbolic truncation q-norm of $q = 0.9$. We consider a validation set having $N_{LHS} = 10^4$ sample points. The input vector is 9 dimensional $\xi = \{\alpha_{V2}, C_{33max0}, C_{33max1}, ad_{31}, \gamma_{1/2}, c_{rk}, n_{high}, n_{low}, \alpha_{L1}\}$ with *uniform* marginal distribution (refer to Table 3.2) assumed on each variable. We use the same experiment (BCF1) to construct our PCE-proxy, which we have used for senstivity analysis (refer section 3.4). We start with $N_s = 50$, which are increased in each iteration step, if the stopping criteria is not met. Figure 3.3 shows the evolution of ϵ_{LOO}^{mean} and ϵ_{emp}^{emp} error with increase in N_s. It can be observed that ϵ_{emp}^{mean} error is a monotonically decreasing function with PCE sample size. Also it often gives an optimistic estimates which leads to overfitting of PCE (refer to section 3.5.2). Therefore, a cross-validation error estimate ϵ_{LOO}^{mean} is also plotted, which estmates how well PCE-proxy predicts on points other than experimental design points. As expected, PCE-proxy constructed with larger training sets estimates more accuartely than the ones constructed with smaller. A sample size of 1000 was sufficient to achieve the desired accuracy $(\Delta\epsilon_{LOO}^{mean})$.

The values of $\Delta\epsilon_{LOO}^{mean}$ estimated in the final iteration step were 4.6×10^{-4} for pressure drop profile and 4.1×10^{-4} for cumulative oil recovery curve. It should also be noted that the total time taken to evaluate simulation model at 1000 sample points was $\approx$ 20 hrs on a system with Intel(R) Xeon(R) CPU E-5 1650 v3 @3.5GHz CPU with 32 GB RAM.

(a) Pressure drop profile   (b) Cumulative oil recovery curve
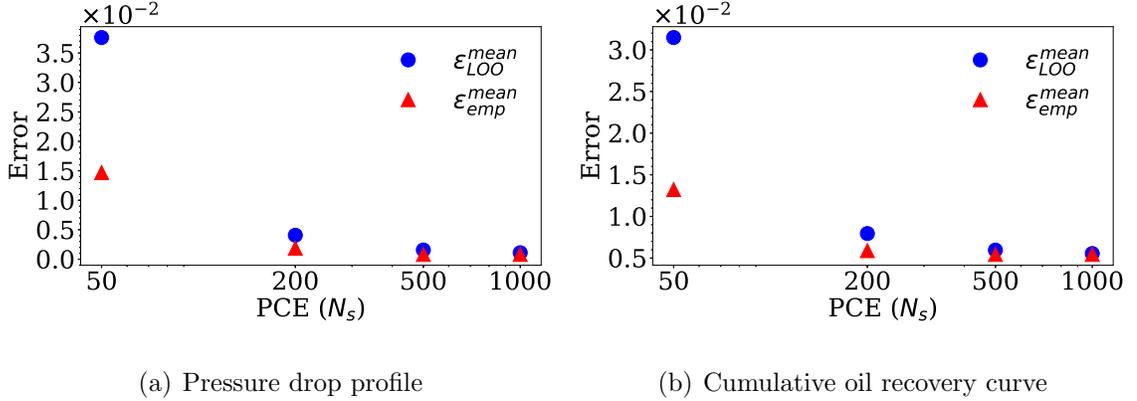
Fig. 3.3. Mean LOO error and mean empirical error for PCE-proxy constructed using $N_s = \{50, 100, 500, 1000\}$ training samples.

We now investigate the predictive ability of PCE-proxy by validating it with simulation results for overall oil recovery factor and $max(\Delta p)$. This can be done either by cross-validation of PCE-proxy response with simulation model response or by comparing their histogram. Figure 3.4 and 3.5 shows the effect of increasing the size of experimental design set ($N_s = \{50, 200, 500, 1000\}$), on predictive capabilities of PCE-proxy. It is clearly evident that increasing N_s results in increased accuracy ($R^2 \to 1$). It is worth noting that even with $N_s = 200$, the PCE-proxy is able to accurately predict the simulation model response for both the overall oil recovery factor and $max(\Delta p)$ with $R^2 \approx 0.99$. For $max(\Delta p)$ response, the coefficient of determination (R^2) improves significantly from 0.965 for $N_s = 50$ to 0.996 for $N_s = 1000$. Figure 3.6 compares the histogram of $max(\Delta p)$ and overall oil recovery factor for PCE-proxy constructed with $N_s = 1000$ samples and simulation model for 10^4 LHS points. There is an excellent agreement between simulation model response and PCE-
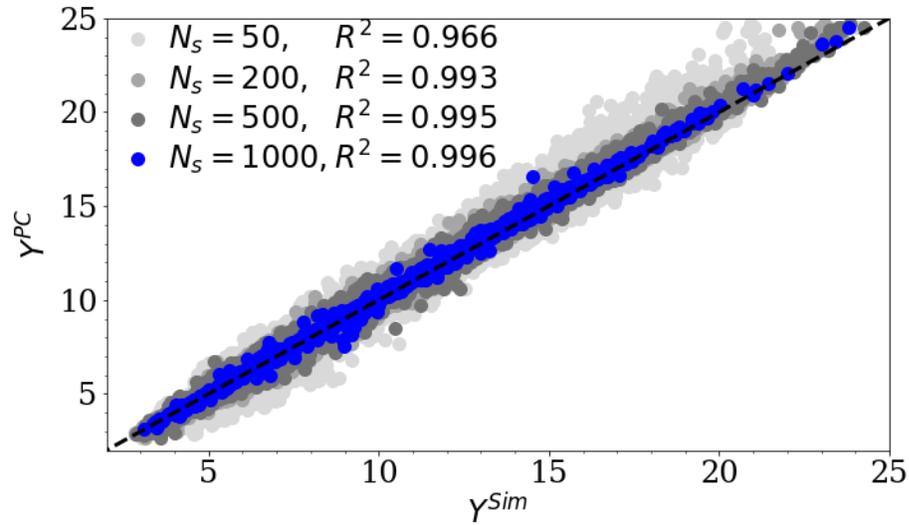
Fig. 3.4. Cross-validation of PCE-proxy (Y^{PC}) constructed with N_s={50,200,500,1000} training sets with simulation model (Y^{sim}) for maximum pressure drop ($max(\Delta p)$).
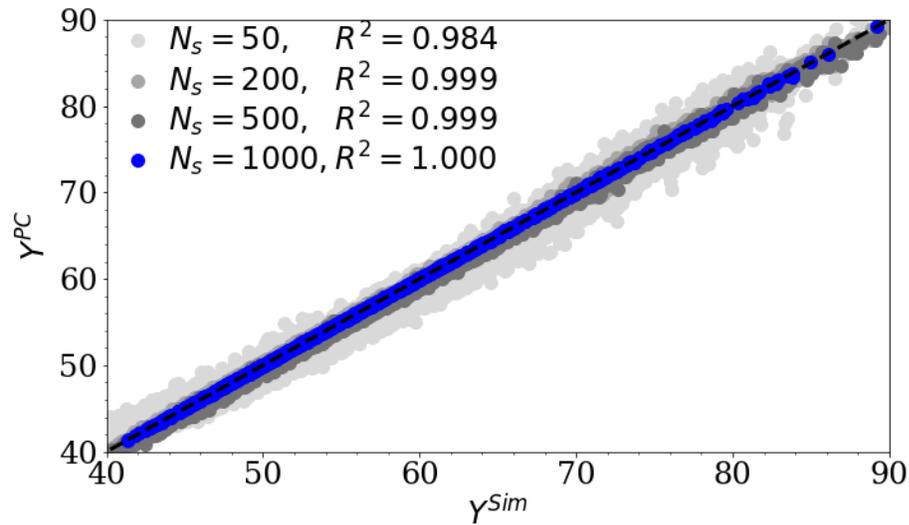
Fig. 3.5. Cross-validation of PCE-proxy (Y^{PC}) constructed with N_s={50,200,500,1000} training sets with simulation model (Y^{sim}) for overall oil recovery factor.

proxy response indicating that the PCE-proxy can be used efficiently in inverse model calibration and history matching.

(a) Overall oil recovery factor   (b) $\max(\Delta p)$
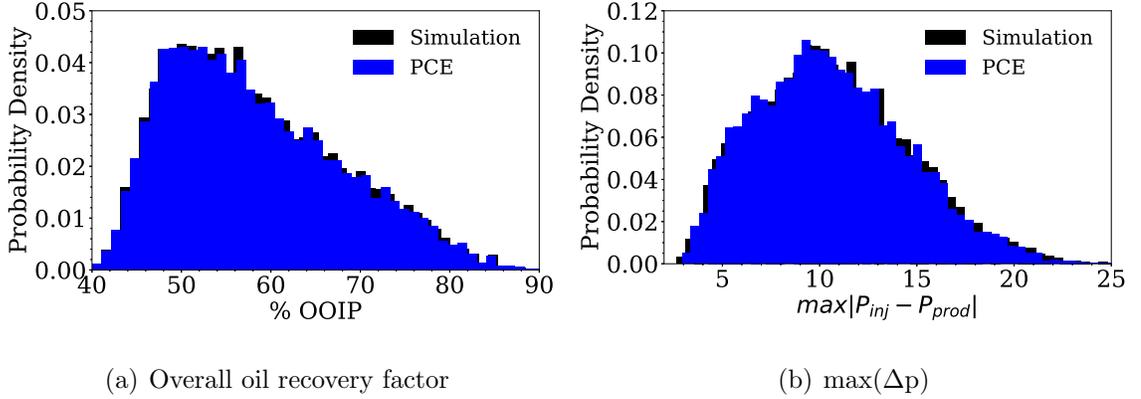
Fig. 3.6. Histogram of PCE-proxy predictions constructed with $N_s=1000$ vs simulation model for 10^4 LHS sample points.

3.5.4   Moments of PCE: Mean and Variance

Moments of stochastic expansion methods such as PCE are available analytically which converge to true output model response. The first two moments of PCE, mean and variance, are functions of only its coefficients. To compute mean (μ_{PCE}), we take an expectation $(\mathbb{E})$ of PCE and then use *inner product* with Ψ_0 as shown below.

$$\mu_{PCE} = \mathbb{E}[Y] = \mathbb{E}_p[P(\xi)] = \mathbb{E}_p[P(\xi) \cdot 1] = \langle P, \Psi_0 \rangle = y_1. \qquad (3.11)$$

Therefore, mean of PCE is given by the first coefficient of constant basis term $\Psi_0=1$. The variance of PCE $(\mathbb{V})$ can be computed using the common formula of variance:

$$\mathbb{V}[Y] = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2.$$

The first term is given by Eq. (3.11) and for the second term we have:

$$\mathbb{E}[Y^2] = \mathbb{E}_p[P^2(\xi)] = \| P \|^2 = \sum_{\alpha=1}^{\infty} y_\alpha^2.$$

Combining everything, we get:

$$\sigma_{PCE}^2 = \mathbb{V}_p[P(\xi)] = \sum_{\alpha=2}^{\infty} y_\alpha^2. \tag{3.12}$$

Therefore, it can been seen that the variance of PCE is given by the sum of square of all its coefficients except the first (y_1).

3.5.5 Moments of stochastic approach (LHS) : Mean and Variance

Consider n independent and identically distributed (iid) samples $(\xi = (\xi_1, \xi_2...\xi_n))$ from any distribution. The mean of the distribution (I_n) can be approximated using *Central Limit Theorem (CLT)*, if variance (V_n) of distribution exists. The *CLT* states that as $n \to \infty$, the empirical average of any distribution tends to a normal distribution with mean:

$$I_n := \frac{1}{n} \sum_{j=1}^{n} P(\xi_j) \to N\left(\mathbb{E}[P(\xi)], \frac{1}{n}\mathbb{V}[P(\xi)]\right), \tag{3.13}$$

The variance of the distribution (V_n) can then be approximated as:

$$V_n = \frac{1}{n} \sum_{j=1}^{n} (P(\xi_j) - I_n)^2. \tag{3.14}$$

Here, we generate $n = 10^4$ of these iid samples using LHS which has a better convergence rate than Monte-Carlo Sampling (MCS) [139].

We then check the accuracy of the PCE-proxy by looking at the relative error in the PCE-proxy estimates. Let $\mathbf{y}_{\text{true}}$ be the *ground truth vector* and $\hat{\mathbf{y}}$ be the *estimated vector*. The relative L_2 error is defined to be:

$$L_2[\mathbf{y}_{\text{true}}, \hat{\mathbf{y}}] = \frac{\| \mathbf{y}_{\text{true}} - \hat{\mathbf{y}} \|_2}{\| \mathbf{y}_{\text{true}} \|_2}, \tag{3.15}$$

where $\| \cdot \|_2$ is the standard Euclidean norm. Here, we consider, ground truth vector as $\mathbf{y}_{\text{true}} = [I_n, V_n]$, where the mean (I_n) or variance (V_n) are obtained from 10^4 LHS simulations (refer to section 3.5.5) and estimated vector as $\hat{\mathbf{y}} = [\mu_{PCE}, \sigma_{PCE}^2]$, where the mean (μ_{PCE}) or variance (σ_{PCE}^2) are obtained from PCE-proxy (refer to section 3.5.4).

(a) Pressure drop profile: Mean

(b) Pressure drop profile: Variance

(c) Cumulative oil recovery curve: Mean

(d) Cumulative oil recovery curve: Variance
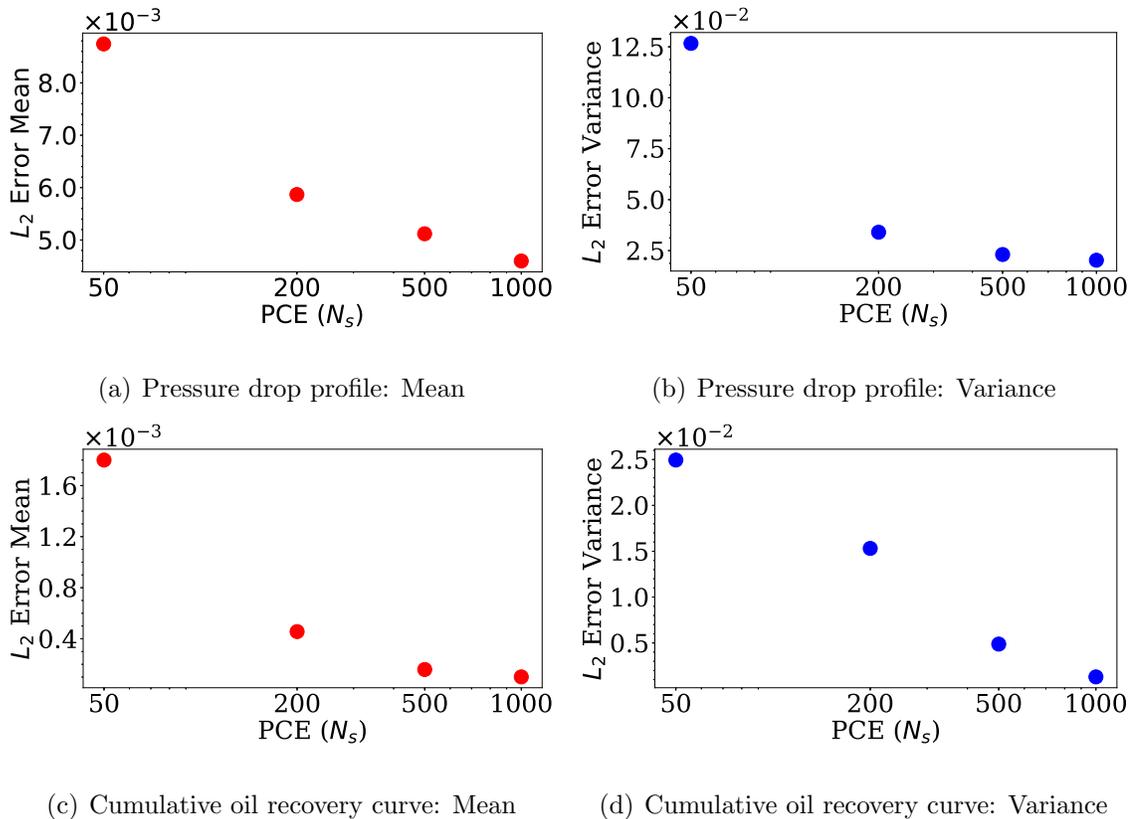
Fig. 3.7. Relative L_2 error in mean and variance of the PCE-proxy output response which converges for all QoIs with increase in N_s. The ground truth is established with 10^4 LHS simulations.

Figure 3.7 shows the relative error in mean and variance estimates of PCE-proxy plotted as a function of N_s. A relative error L_2 of 2.5×10^{-2} is considered to be sufficient for convergence of both the mean and variance plot. It can be seen that PCE-proxy approximates mean within this accuracy even for $N_s = 50$ for both pressure drop profile and cumulative oil recovery curve ($L_2 \sim O(10^{-3})$). Even for variance, it takes only about $N_s = 50$ for PCE-proxy to achieve this high accuracy for cumulative oil recovery curve. However, for pressure drop profile it takes about $N_s = 1000$ to achieve this convergence criteria. This is because of large variance or uncertainty associated with pressure drop profile after surfactant-polymer flooding (after 1.45

PV) as seen in figure 3.9. A more detailed discussion on this is provided in section 3.8 where we use PCE-Kriging technique to quantify this uncertainty.

In figures 3.8-3.9, we plot mean (μ_{PCE}) and standard deviation (σ_{PCE}) for our QoIs to visualize how fast PCE converges to the same solution statistics (I_n and σ_{LHS}) as estimated by 10^4 MC simulations performed using LHS sampling. As it can be seen, even with $N_s = 50$, the mean and stand deviation are estimated very accurately (refer to relative L_2 error reported in figure 3.7), which shows the great potential PCE offers for computational speed up. It is worth noting that it takes only about few seconds to evaluate PCE at 10^4 experimental design points. This is very important for inverse optimization problem which usually requires few thousands of model evaluations in each iteration step (refer to section 3.6).

(a) Pressure drop

(b) Cumulative oil recovery
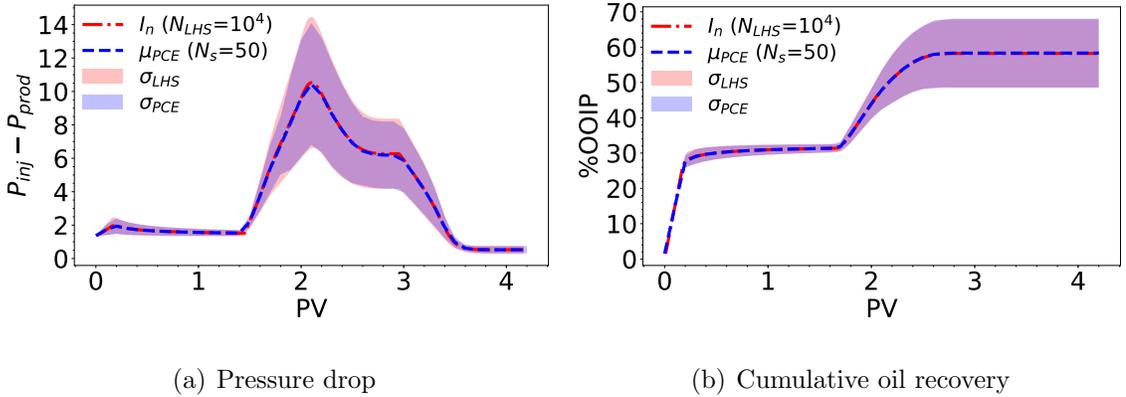
Fig. 3.8. Comparison of mean and standard deviation of PCE-proxy constructed with $N_s = 50$ with 10^4 LHS evaluations for pressure drop profile and cumulative oil recovery curve.

3.6 Inverse Model Calibration Problem

Let Ξ be an n-dimensional space of uncertain model parameters (here $n = 9$). Ξ contains all the prior possible combinations of unknown parameter values ξ. Let the experimental observations be denoted with y. Let P be a forward operator of the

(a) Pressure drop

(b) Cumulative oil recovery
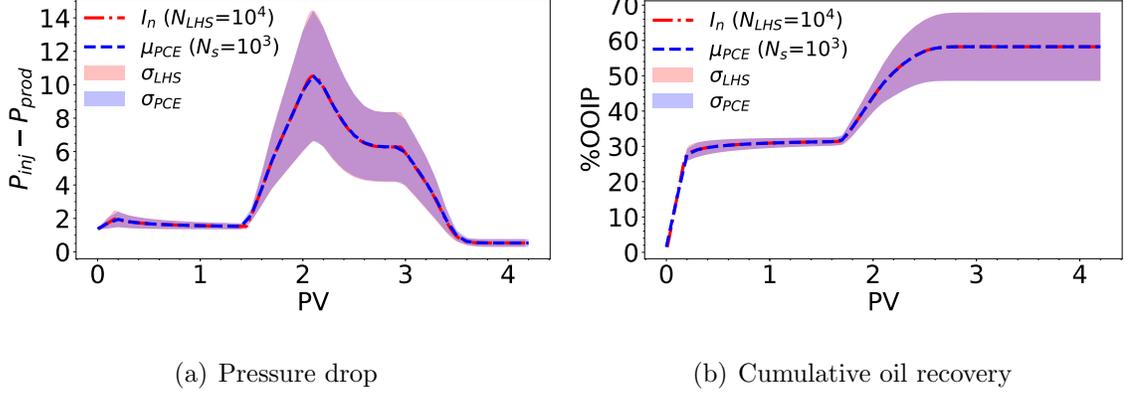
Fig. 3.9. Comparison of mean and standard deviation of PCE-proxy constructed with $N_s = 1000$ with 10^4 LHS evaluations for pressure drop profile and cumulative oil recovery curve.

PCE-model, which is a polynomial function that maps the space of Ξ to the space of y's (Y). These uncertain model parameters can be adjusted to achieve an acceptable fit between the model and observations. The *inverse problem*, also known as the *model calibration* problem is to find the **best** $\bar{\xi}_{min} \in \Xi$ such that $P(\bar{\xi}_{min}) \approx y$.

We define a *loss metric* $\ell(P(\xi), y)$, to quantify how closely function $P(\xi)$ approximate our QoIs, y, viz. cumulative oil recovery curve and pressure drop profile (refer to section 3.3). Therefore, we can setup a multi-objective optimization problem with a separate loss metric defined for each one of these objectives. Here, we use a *square loss* metric which is given as:

$$\ell(P(\xi), y) = \frac{1}{2} \parallel P(\xi) - y \parallel_2^2 = \frac{1}{2} \sum_{j=1}^{k} (P_j(\xi) - y_j)^2, \tag{3.16}$$

Therefore, our optimization or model calibration problem is to find $\bar{\xi}_{min} \in \Xi$ such that it **minimizes** $Q = [q_1, q_2]$, where

$$q_s = \min_{\xi \in \Xi} \frac{1}{2} \parallel P(\xi)_s - y_s \parallel_2^2. \tag{3.17}$$

Here, s=1 is cumulative oil recovery curve and s=2 is pressure drop profile.

3.6.1 Genetic Algorithm based Optimization

We pose the above inverse problem as an optimization problem (refer to section 3.6) and use *Genetic Algorithm* (GA) to solve it. GA is based on the principle of survival of fittest (refer to Algorithm 3) that mimics natural selection. GA was first proposed by Holland in 1975 [147] and later developed by Goldberg in 1989 [148]. Since then, GA has been extensively applied to a variety of applications such as uncertainty quantification [149, 150], reservoir history matching [151, 152], optimization [153, 154] or model calibration [155, 156].

The procedure starts with a generation of initial feasible random population, with each member of this population known as *chromosome*. Fitness of each chromosome is evaluated using an objective function called as *fitness function*. The fittest chromosomes among them called as *parents* are evolved through *mutation* and alteration (*crossover*) in subsequent iterations called as *generations* to reproduce *off-springs*. If these off-springs perform better, the original chromosomes in the population are replaced with them. This process is repeated for many generations till the solution converges to the best set of chromosomes which is considered as the optimized set based on natural selection.

3.6.2 Parent Selection

The randomly generated population ($pop_{size} = 120$) is evolved by selecting only those individuals from the population who has a better probability of surviving, finding mate and reproducing fitter off-springs. This is achieved by defining a *selection function* where these individuals are competed against each other based on their fitness functions. Most common types of selection functions found in literature are tournament selection [157], roulette wheel selection [158], steady state selection [159], etc.

Here, we use a tournament selection function, where $T_s = 4$ individuals are randomly chosen from the current population pool with probability $P_s = 1$. The fitness

of these T_s individuals is compared and the one with the best fitness is selected. This process is repeated till the population pool is complete. T_s is called tournament size and can be used to control the selection pressure.

3.6.3 Reproduction

Off-springs are reproduced when two parents *mate*. At the molecular level, mating involves exchange of genetic information from both the parents. This is accomplished by splitting the chromosomes of both parents at a randomly selected point (single-point crossover) or at multiple points (k-point crossover, uniform crossover) and crossing them over. This crossover operation is carried with a crossover probability P_c of 0.8 meaning 80% of the new population is reproduced by crossover operation. Crossover improves the search of optimum solution by exploiting a larger parameter space. However, only crossover results in premature convergence because of over-exploitation of search space often resulting in a suboptimal convergence. Therefore, these off-springs are mutated with a small mutation probability P_m by randomly altering their genes. This ensures that genetic diversity is maintained among off-springs and results in better exploration of the parameter search space. But a high mutation probability could mean that the algorithm is randomly guessing the optimum solution and hence, it may suffer from slow-convergence. We use 0.01 as mutation probability P_m. More information on these operators can be found in [148, 160, 161].

3.6.4 Fitness function

We use Eq. (3.17) as fitness function, where the objective is to minimize the square of difference between observations and PCE-proxy. We observe that these two objectives are conflicting in nature with respect to the uncertain input vector x. This means that a single optimum solution is not possible and a decrease in one objective function results in increase in another objective function. Therefore, a concept of Pareto dominance is used to define the trade-off between these multi-

objective functions [162]. In Pareto dominance, we seek to find a set of all equally dominant solutions called Pareto optimal that provides an optimum trade-off. ξ^* is called as Pareto optimal S [163], if for all ξ, $S_s(\xi^*) \geq S_s(\xi)$ for $s = 1, 2$ and $S_s(\xi^*) > S_s(\xi)$ for at least one s. This implies that, for any feasible ξ, decrease in one or more objective functions will result in increase in at least one objective function. Therefore, we are expected to obtain a set of Pareto optimal solutions Ξ^*, which is a subset of Ξ. The image of all $\xi^* \in \Xi^*$ is called Pareto front, representing this trade-off. Here, we consider Pareto optimal to be the optimized solution set $(\bar{\xi}_{min})$ (refer to section 3.6) i.e. $\bar{\xi}_{min} = \xi^*$.

Figure 3.10 shows the Pareto front after $N_{gen} = \{1, 10, 50, 100\}$ generations of Genetic Optimization Algorithm. The algorithm is stopped after N_{gen}^{end} generations (in line 14 of Algorithm 3). The trade-off between the error for cumulative oil recovery curve (horizontal-axis) and pressure drop profile (vertical-axis) can be clearly seen. It can also be concluded that GA does a great job of minimizing both of the objective functions since the Pareto front shifts towards the origin with increasing generations.
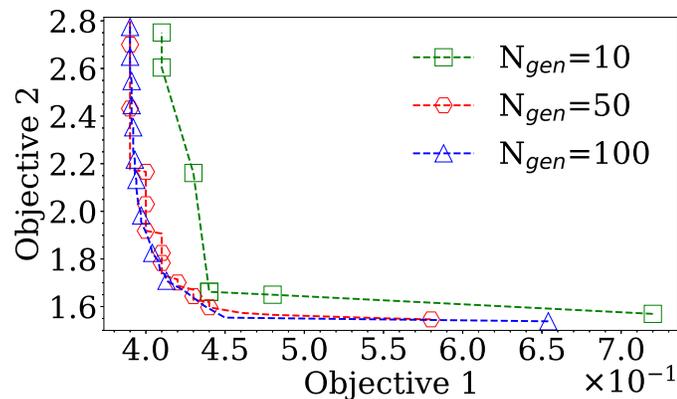
Fig. 3.10. Pareto front showing the trade-off between the two objective functions.

3.7   Final model calibration results

The model calibration algorithm (refer to Appendix A) starts with minimizing both of the objective functions (refer to Eq. (3.17)). This multi-objective optimization ensures a simultaneous reduction of the misfit between the observed data and PCE-proxy response of our QoIs for BCF1. Recall that the PCE-proxy is constructed using injection sequence and slug design of BCF1. The stopping criteria used for this stage is $n_{cross} = 50$, where after every n_{cross} generations the predicted optimum by GA is cross-validated with BCF2 and BCF3 (in Algorithm 3 line 6). This is done by calculating an average relative L_2 misfit error (σ_{loss}) for BCF2 and BCF3 simulation model for optimal points on the Pareto front. This cross-validation stage ensures that the minima predicted with GA is actually a *global minimum* by predicting the response of calibrated models on two different independent data sets BCF2 and BCF3. The model calibration algorithm is stopped, when maximum generation limit exceeds N_{gen}^{max}=100 or σ_{loss} is reduced below an acceptable value.

Figures 3.11-3.13 show the model calibration algorithm in working, where the simulation model is evaluated at all the points on the Pareto front after $N_{gen} = 1$ and $N_{gen} = 100$ generations and is compared with experimental observations. Note that, here we use simulation model as the proxy model is constructed only for BCF1 but not for BCF2 and BCF3. The green shaded area represents the miss-fit between observations and simulation model predictions after $N_{gen} = 1$. This miss-fit is clearly seen to be reduced (red shaded area) after $N_{gen} = 100$ generations of GA optimization for pressure drop, cumulative oil recovery curve as well as for oil-cut curve. The same model parameters can be seen to accurately predict the QoIs for BCF2 and BCF3 without any further tuning indicating convergence to a global optimum. Figure 3.14 compares each of these calibrated models separately before and after optimization. The original larger parameter space from uniform distribution (light shaded area) is clearly reduced and converged to a much smaller parameter space represented by dark shaded areas. Intersted readers can read more about these models in [27, 65, 128].

(a) Pressure drop (b) Cumulative oil recovery (c) Oilcut
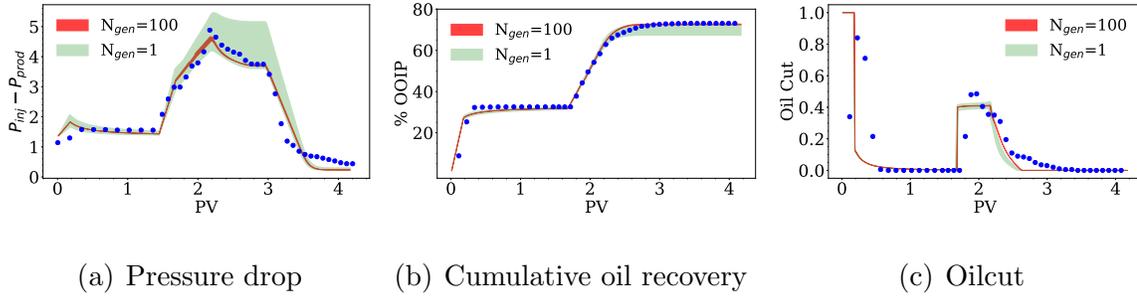
Fig. 3.11. History matching for BCF1 where the response is plotted for the points on the pareto front after $N_{gen} = 10$ and $N_{gen} = 100$ for a) pressure drop b) cumulative oil recovery c) oilcut curve. The parameter space after $N_{gen} = 10$ can be seen to be reduced and converged well to the experimental data after $N_{gen} = 100$.

(a) Pressure drop (b) Cumulative oil recovery (c) Oilcut
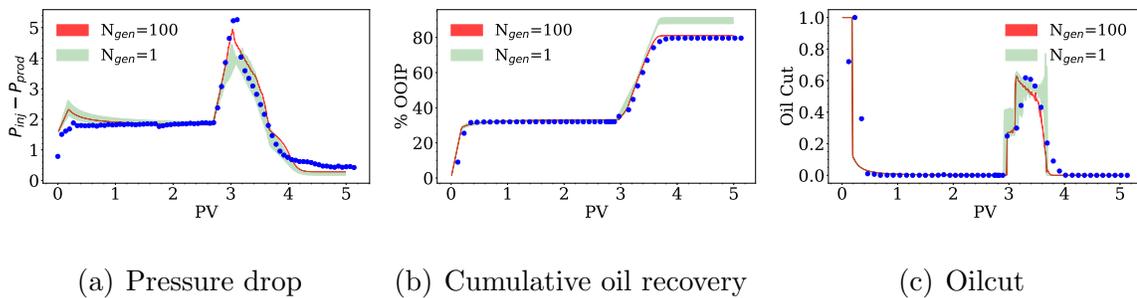
Fig. 3.12. Validation for BCF2 where the response is plotted for the points on the pareto front after $N_{gen} = 10$ and $N_{gen} = 100$ for a) pressure drop b) cumulative oil recovery c) oilcut curve. The parameter space after $N_{gen} = 10$ can be seen to be reduced and converged well to the experimental data after $N_{gen} = 100$

(a) Pressure drop   (b) Cumulative oil recovery   (c) Oilcut
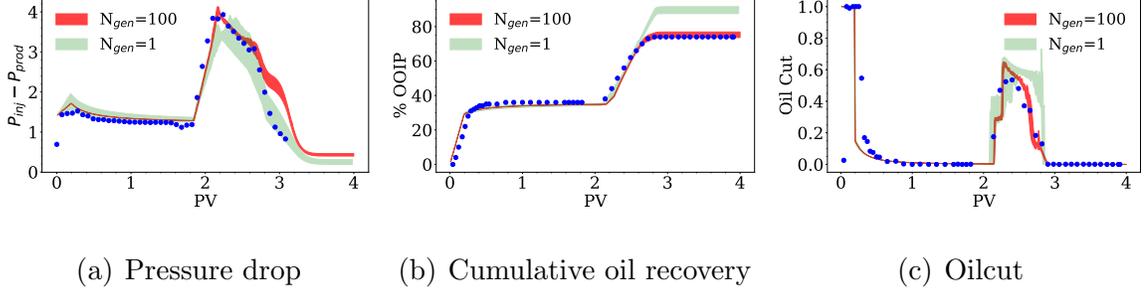
Fig. 3.13. Validation for BCF3 where the response is plotted for the points on the pareto front after $N_{gen} = 10$ and $N_{gen} = 100$ for a) pressure drop b) cumulative oil recovery c) oilcut curve. The parameter space after $N_{gen} = 10$ can be seen to be reduced and converged well to the experimental data after $N_{gen} = 100$

3.8   Uncertainty quantification in PCE-proxy

In this section, we explain the formulation of PCE-Kriging method to quantify epistemic uncertainty in PCE-proxy model [164, 165]. Kriging is based on stochastic interpolation algorithm which uses weighted average method to interpolate value of functions at unobserved points. Kriging is also known as Gaussian process regression. It approximates model output as:

$$Y^{PCEK} = f(\xi) \approx \beta^T P(\xi) + \sigma^2 Z(\xi, w). \tag{3.18}$$

This Gaussian regression process has two main ingredients i) a mean function $\beta^T P(\xi)$ and ii) a covariance function which is characterized by variance σ^2 and a zero mean and unit variance stationary Gaussian $Z(\xi, w)$. When the mean is approximated using a PCE trend, the Kriging method is known as PCE-Kriging or PCE-K in short. Here, we adopt the framework of PCE-K from UQLab [133] which is based on [120]. Therefore, substituting Eq. (3.4) in Eq. (3.18), we get:

$$Y^{PCEK} \approx \sum_{\alpha \in A} y_\alpha \Psi_\alpha(\xi) + \sigma^2 Z(\xi, w). \tag{3.19}$$

(a) Relative perm. at high Ca.

(b) Relative perm. at low Ca.

(c) Micro-emulsion Viscosity

(d) Residual resistance factor
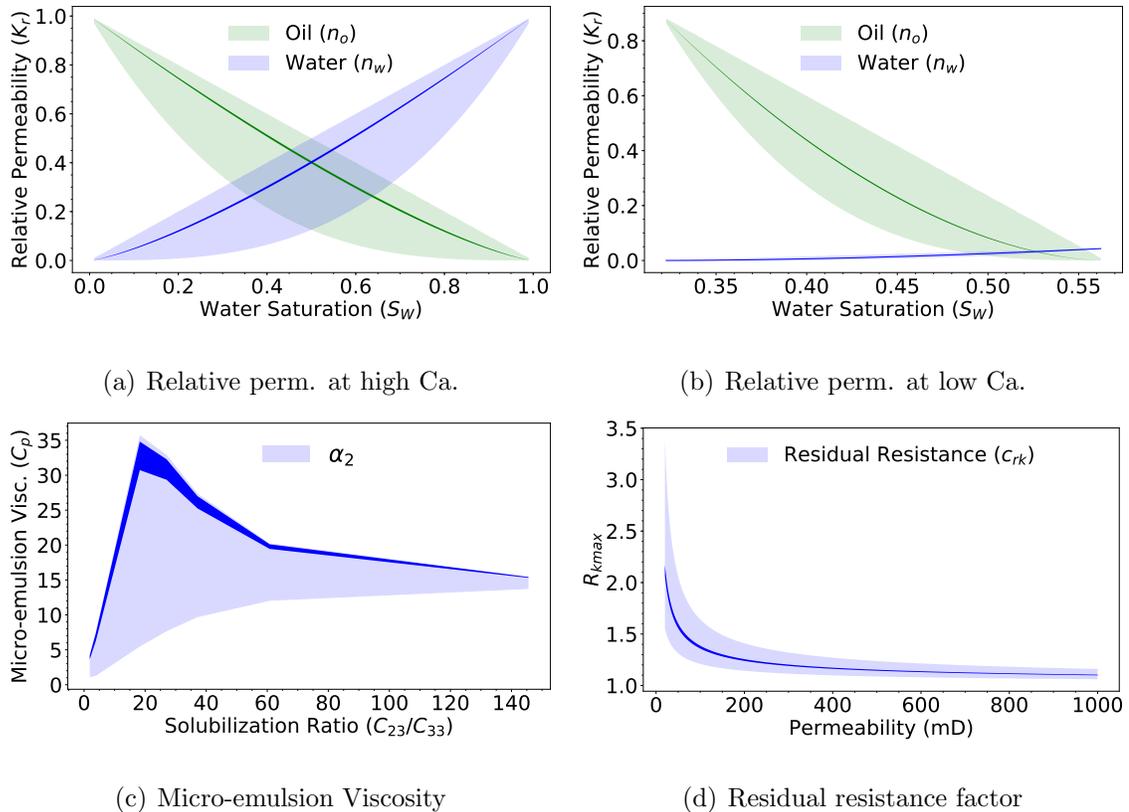
Fig. 3.14. Final calibrated models using a Genetic Algorithm: light shaded area represents the parameter space before calibration, while dark shaded area represents the optimized parameter space after calibration

This is a *universal* Kriging model, where the mean is obtained from weighted sum of orthonormal polynomial basis. Therefore in PCE-K, PCE captures the global behavior of the trend while the Gaussian process captures the local variability. Eq. (3.19) can be solved using *sequential PCE-K* approach as described in [120]. Stationary Gaussian $Z(\xi, w)$ is characterized by its correlation function $\mathbf{R}(\xi, \xi'; \theta)$, which describes the nature of function space. In more simpler words, the choice of correlation function depends on how similar are observations $f(\xi)$ and new points $f(\xi')$ in the function space depending on the distances between them. We use *Squared Exponential* as correlation function or kernel R which is given as:

$$R(\xi, \xi', \theta) = exp\left(- \left(\frac{\xi - \xi'}{\theta}\right)^2\right), \tag{3.20}$$

where θ is the correlation length, which is generally not known because of lack of observations. However, an estimate for θ can be calculated by setting up a Maximum Likelihood optimization problem which reads [166]:

$$\theta = arg\ min\left(\frac{1}{2}log(det(R)) + \frac{N_s}{2}log(2\pi\sigma^2) + \frac{N_s}{2}\right). \tag{3.21}$$

The idea is to find θ that maximizes the likelihood of observations:

$$Y = \{f(\xi_1), f(\xi_1), ...f(\xi_{N_s}^T)\}.$$

The optimal value of θ can then be used to obtain other hyperparameters of Kriging process (σ^2, coefficients β^T).

Figure 3.15 shows uncertainty quantification in the proxy model using PCE-K, which is trained using $N_s = 200$ LHS samples. Here, the mean of PCE-K proxy model is also compared with PCE predictions (refer to section 3.5). Eq. (3.21) is solved using a local optimization algorithm called L-BFGS Hessian. We use the same input parameter space and design settings as used in PCE construction step (refer to section 3.5).

(a) Pressure drop   (b) Cumulative oil recovery   (c) Oilcut
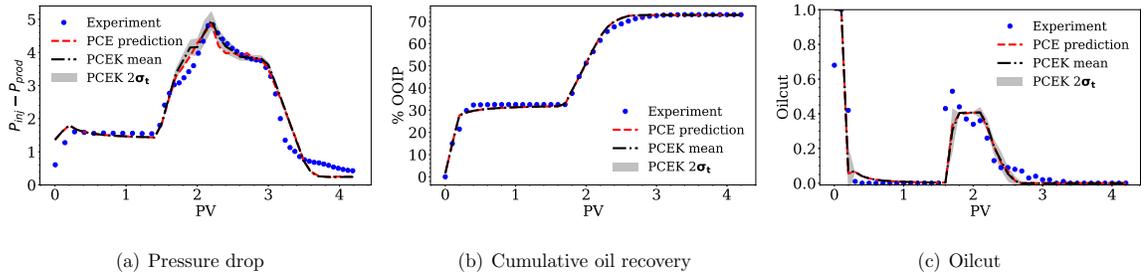
Fig. 3.15. Uncertainty quantification in proxy model using PCE-Kriging. The uncertainty can be observed to be high during SP and P flooding because of the large uncertain parameter space describing their mechanism.

It is interesting to see that most of the epistemic uncertainty ($2\sigma_t$) in PCE-K model comes from surfactant-polymer injection (after 1.45 PV). As explained in section 3.3, surfactant lowers the interfacial tension (IFT) between the oleic phase and aqueous phase and can lead to the formation of a new phase called as micro-emulsion. Modeling such a process requires many physical models or parameters (Table 3.1) and hence involves a considerable uncertainty. This uncertainty is captured very well by a PCE-K proxy model. On the contrary, there is little to no uncertainty during initial water flooding (IWF) (0-1.45 PV) and extended water flooding (EWF) (2.96-4.3 PV) phases, signifying that PCE mean captures the trend nicely.

3.9 Conclusion

Even though there exits an exhaustive literature on history matching processes, there isn't a single algorithm that works best for all problems. Rather, the choice of algorithm depends solely on the problem at hand. Therefore, any new algorithm that can overcome some of the drawbacks of already existing ones, finds a great interest in the community [167]. Here, we propose an algorithm to construct a very fast and an efficient emulator for history matching a surfactant-polymer (SP) flood process by exploiting the characteristics of model response through sensitivity analysis and proxy modeling.

At first, the uncertain model parameters necessary to model SP flood process are defined (19 variables in total). The high-dimensionality of this stochastic parameter space is reduced by implementing a Sobol SA. This low-dimensional parameter space (9 variables in total) obtained from SA is then used to construct an efficient proxy model based on Polynomial Chaos Expansion (PCE-proxy). PCE-proxy approximates the highly non-linear response of simulation model and guarantees convergence in distribution of output quantities of interest. The predictive capabilities of PCE-proxy is cross-validated with simulation model for experimental design size of $N_s = \{50, 200, 500, 1000\}$ training samples for overall oil recovery factor and max Δp. The

analytical moments of PCE-proxy viz. mean and variance are also compared with statistical moments derived from benchmark simulations of 10^4 LHS samples and it is shown that even experimental design set of $N_s = 50$ points exhibit an acceptable convergence. Then, an inverse optimization problem is proposed to minimize the square loss metric or miss-fit between PCE-proxy and observations. This multi-objective optimization problem is solved using a GA. Many local minima exits for this optimization problem because of the non-linearity of the model response. Therefore, GA is expected to perform better as it starts with multiple points in the search space. Use of PCE-proxy in this optimization stage facilitates an accelerated approach to determine an optimized parameter space and improve the quality of history matched results.

The proposed framework has been successfully applied to history match three different core-flood experiments for all quantities of interests such as pressure drop and cumulative oil recovery curve. Finally, uncertainty arising in the proxy model itself is quantified by constructing a PCE-Kriging proxy model. This PCE-K proxy model uses PCE as basis to capture global variability or trend and a Gaussian regression process to capture the local variability thus quantifying uncertainty in the proxy model construction stage.

The main disadvantage of the presented framework is for higher order approximations where an increase in PCE degree directly reflects to a drastic increase in PCE terms which requires a large number of simulation runs for their computation. However, recently many new methods have been proposed which exploit the sparsity-of-effects principle to detect the main-effects and lower-order interaction terms that greatly governs the model response, neglecting many terms in PCE-proxy expansion [168–170]. One potential extension of the presented framework could be in the Bayesian context where cheap PCE-proxy replaces the full-forward model in likelihood function to obtain an analytical expression of posterior or to directly sample from the posterior using MCMC methods [68]. We discuss this approach in the next Chapter 4.

In all, the methodology presented here can be extended to a class of more complex model calibration problems, such as history matching of fractured reservoirs [171, 172], history matching of alkaline-surfactant-polymer (ASP) flooding and optimization [173] or CO_2-EOR and storage [174, 175]. The framework presented can also be applied to solve problems outside the realm of reservoir engineering as it provides a way to validate accuracy of surrogate model, quantify uncertainty in the surrogate model and provides a probabilistic measure of model output on stochastic inputs making model calibration task extremely efficient.

4. BAYESIAN MODEL CALIBRATION AND OPTIMIZATION

A version of this chapter is in preparation for submission as:

Pratik Naik, Piyush Pandita, Soroush Aramideh, Ilias Bilionis, & Arezoo Ardekani, Bayesian model calibration and optimization of surfactant-polymer flooding. *Computers & Geosciences*.

4.1 Abstract

The underlying physical models governing surfactant-polymer (SP) flooding process are subject to parametric uncertainties, accurate quantification of which is crucial for improved decision making. Moreover, history matching of SP flooding is an ill-posed problem, typically characterized by a multi-modal posterior distribution of these model parameters. This chapter presents a systematic approach for Bayesian history matching and uncertainty quantification in the model calibration stage of SP flooding using coreflood experimental data. The approach is as follows. First, we construct a surrogate of the computationally expensive physics-based model using a polynomial chaos expansion (PCE-proxy). Second, we formulate a Bayesian calibration problem for inferring the model parameters from a single coreflood experiment that measures pressure drop and oilcut profiles. Third, we solve the Bayesian calibration problem by sampling directly from the posterior using Markov chain Monte Carlo (MCMC). We validate the calibrated parameters by successfully predicting the result of two other coreflood experiments. Then, we extend this framework to stochastic multi-objective optimization of injection slug size design under uncertainties in model parameters (captured by the posterior of Bayesian calibration problem) and oil price (modeled as a geometric random walk with constant drift and volatility). To identify

the Pareto frontier of the stochastic multi-objective optimization problem, we employ a variant of Bayesian global optimization (BGO), a class of algorithms capable of optimizing black-box, gradient free, computationally expensive functions. In particular, we use the extended expected improvement over the dominated hypervolume to sequentially select simulations that seek to reveal the Pareto frontier. An addendum of the implemented BGO is that it quantifies the epistemic uncertainty about the Pareto frontier as induced by the limited number of simulations used to construct it.

4.2 Introduction

Surfactant-polymer (SP) flooding for chemically enhanced oil recovery (cEOR) has been extensively studied over the past few decades [5, 45, 46]. In SP flooding, injection of polymer to the oil-water system hinders the flow instabilities and results in an improved sweeping efficiency [28, 47] whereas injection of surfactant results in reduced interfacial tension (IFT) which leads to mobilization of the trapped oil [48, 49]. Conventionally, laboratory scale coreflood experiments are preformed to gain better understanding of the characteristics of surfactant-polymer flooding processes. However, major technical challenges occur while transferring the SP flood technology from laboratory scale to the reservoir scale [50]. This is mainly because of uncertainties associated with the local reservoir properties such as porosity, permeability and saturations [176, 177] or uncertainties in the parameters governing the surfactant and polymer flooding process which is the focus of this chapter [178]. Therefore, stochastic modeling techniques are often used to explicitly capture the variability in these parameters and their associated uncertainty [50].

In the literature, various stochastic modeling techniques have been successfully applied to the reservoir history matching problem such as gradual deformation method (GDM) [179], ensemble Kalman filter (EnKF) [180, 181] or randomized maximum likelihood (RML) method [182]. GDM builds a stochastic process by generating perturbations of prior model sample with an aim that the perturbed realization grad-

ually matches well with the observed data [183]. This perturbation is achieved using trigonometric functions of Gaussian random variables [184]. Though the perturbation method preserves the prior model statistics, Caers [185] showed that when the problem is framed in Bayesian context, the sample drawn is not guaranteed to be from the true posterior distribution. EnKF is a sequential assimilation method where an ensemble of model realizations is updated based on deviation of the model response from the measured data [82]. Though this method has been extensively studied for reservoir history matching, it suffers from various limitations in case of non-Gaussian prior information, non-linearity of the forward model or scaling of the model to the field scale application [186]. Also, EnKF, if it is not combined with suitable parameterization method, results in extreme non-physical values for the model parameters [187].

To quantify the epistemic uncertainty in the model parameters, one has to pose the model calibration problem in a Bayesian context. Bayesian inference is attractive because it reduces the problem of history matching to just sampling the posterior probability density function (PDF) of the model parameters conditioned on the field data by building a probabilistic model of the process. RML is one such method that can be used to sample this posterior PDF of the model parameters [188]. However, RML has shown poor performance when the posterior is non-Gaussian or the forward model is non-linear, which is the case in SP flooding [189]. Another appealing alternative to sample from the posterior is to use rigorous iterative sampling techniques such as Markov chain Monte Carlo (MCMC). Even though MCMC is computationally expensive, by the law of large numbers, it is theoretically guaranteed to provide samples from the true posterior as the number of iterations goes to infinity.

MCMC has a long history in reservoir history matching applications [190]. Recently, many new advances have been made to the traditional MCMC technique such as hybrid MCMC [191], parallel interacting MCMC [192], population MCMC [193], combining EnKF with MCMC [189] or two-stage MCMC [194]. The main focus of all these improvements has been to address the prohibitively expensive cost of MCMC as evaluating full scale reservoir model even for few thousands of proposed MCMC sam-

ples is computationally unfeasible. In this chapter, we address this issue by constructing a Polynomial Chaos Expansion-based proxy model (PCE-proxy) which replaces the original expensive physics-based model in the likelihood function of Bayesian inference. This PCE-proxy approximates the non-linear output response of simulation model and facilitates a significant reduction in computational time [195]. Proxy-based Bayesian history matching has been studied extensively in recent years using artificial neural networks (ANN) [196], K Nearest Neighbouring (KNN) [197,198], response surface methodology (RSM) using polynomial regression design of experiments [199], Gaussian process regression [200–202], PCE-proxy [203,204]. PCE-proxy is the choice of surrogate model here because of the low-dimensionality of the calibrated parameters (9 parameters are considered), it can be built easily in a non-intrusive manner by treating the forward simulation model as a black-box, and, for smooth functions, it converges exponentially fast as the number of model simulations and polynomial PCE degree increases [195].

In this chapter, we present an assisted history matching framework using PCE-proxy based MCMC strategy to calibrate models for SP flooding process in Berea corefloods (BCF). We pose the problem in Bayesian context where at first, the *prior* knowledge of various physical models imitating the process of SP flooding is obtained by conducting separate experiments which are referred as *support experiments*. Then, the uncertain parameter space of SP flood process is represented via appropriate orthonormal basis of PCE-proxy model. This PCE-proxy is used to approximate realizations of quantities of interest (QoIs) such as pressure drop and oilcut profile for the proposed MCMC samples. We use this framework to estimate the posterior probability density function of the uncertain model parameters by training the Bayesian model on a single coreflood experiment. We then validate that the trained model accurately predicts our QoIs for two other coreflood experiments with the same posterior. To the best of our knowledge, this is the first study that implements PCE-proxy based Bayesian inference algorithm for history matching of SP flooding.

We then extend this framework to optimize the injection sequence of SP flooding process under uncertainties in these calibrated model parameters by posing a two-objectives optimization problem where these competing objectives are determined by identifying a set of optimal solutions known as the Pareto frontier (PF). The noisy simulator outputs are not mathematically tractable and hence we need to resort to techniques other than the conventional heuristic based optimization techniques [205, 206]. Bayesian global optimization (BGO) [207–210] has been successful in dealing with the above challenges in solving single-objective optimization problems and has recently been extended to multi-objective optimization. BGO models the objectives via Gaussian process surrogates [211] and uses epistemic uncertainty to define an information acquisition function (IAF) that quantifies the merit of evaluating the objective at untried designs. The expensive objective is evaluated at the design corresponding to the maximum value of the IAF, and the latest observation is used to update the surrogate. This iterative process continues until a stopping criterion is met. The most commonly used IAF is the Expected Improvement [207, 208], which extends to the Expected Improvement over the dominated Hyper Volume (EIHV) [212, 213] when solving multi-objective optimization problems. Unfortunately, the current versions of EIHV are unable to deal with parametric uncertainties or stochasticity in the objectives. In this work, we employ an extension of this known as extended EIHV (EEIHV) [214] which filters out the noise due to parametric uncertainties. An addendum of the probabilistic nature of this methodology is that it enables us to characterize our confidence about the predicted PF. We study this optimization problem for two cases of positive and negative drifts in the oil price time series by incorporating a random walk stochastic model for oil price with an aim to capture the influence of volatile oil price on injection strategy.

The chapter is organized as follow. Section 4.3 presents the framework incorporated for Bayesian history matching and optimization along with the details of the BCF experiments. In section 4.4, we briefly discuss the uncertainties present in modeling a typical SP coreflood process. Then in section 4.5, we provide the details of

PCE-proxy model and its validation. Section 4.6 gives the theory of model calibration using Bayesian inference. Finally, in section 4.7, we pose a multi-objective optimization problem using BGO to design the optimum injection sequence for SP flooding under uncertainties in model parameters and oil price.

4.3 Framework for Bayesian history matching

The framework proposed (refer to figure 4.1) involves the sequential execution of initial parameterization, dimensionality reduction, proxy construction, Bayesian model calibration and multi-objective optimization stages. We start by performing laboratory scale experiments to investigate the mechanism of oil recovery by SP flooding. We carry out these experiments with an aim to study the effects of chemical composition specifically salinity and concentration on the overall oil recovery. Therefore, these experiments are performed at different salinity, slug sizes and polymer concentrations as given in Table 2.5. More details on the experimental setup can be found in [65]. We carry out the numerical modeling of SP flooding process with the reservoir simulator UTCHEM v9.82 [27].

An initial model calibration of SP flooding is performed by dividing the entire flood process into three separate sub-processes 1) surfactant flooding 2) polymer flooding and 3) displacement process. The physical models describing each of these sub-processes are calibrated by performing additional support experiments as described in [65]. Table 4.1 shows the different models describing each of these sub-processes along with their parameters. Interested readers may refer to [27] to read more about these models.

Then, the resulting high-dimensional input parameter space (19 stochastic parameters in total) is reduced to 9 parameters by performing a Sobol sensitivity analysis (SA) with output QoIs i.e. maximum pressure drop and overall oil recovery factor. The details of SA are illustrated in our previous work [195] and are also summarized in Table 4.1.

Table 4.1.

Models corresponding to each sub-process and their model parameters.

Sub-process	Model Name	Model Parameters	Sensitive Par.	Min	Max	Source
Surfactant flooding	Micro-emulsion Viscosity	$\alpha_{V1}, \alpha_{V2}, \alpha_{V3}, \alpha_{V4}, \alpha_{V5}$	α_{V2}	2.5	3.0	[4, 127–129]
	Phase Behavior	$C_{33\max 0}, C_{33\max 1}, C_{33\max 2}$	$C_{33\max 0}$	0.02	0.035	Experiment and [60]
			$C_{33\max 1}$	0.01	0.02	Experiment
	Adsorption	ad_{31}, ad_{32}, b_{3d}	ad_{31}	0.1	1.0	Experiment
Polymer flooding	Viscosity and Shear Rate	$P_\alpha, \dot{\gamma}_{1/2}$	$\dot{\gamma}_{1/2}$	50	70	Experiment
	Permeability Reduction	c_{rk}, b_{rk}	c_{rk}	0.06	0.08	Experiment
Displacement process	Relative Perm. (Corey Model)	n_{high}, n_{low}	-	1.5	2.5	[130]
	Dispersivity	$\alpha_{L1}, \alpha_{L2}, \alpha_{L3}$	α_{L1}	0.008	0.012	Experiment and [131]
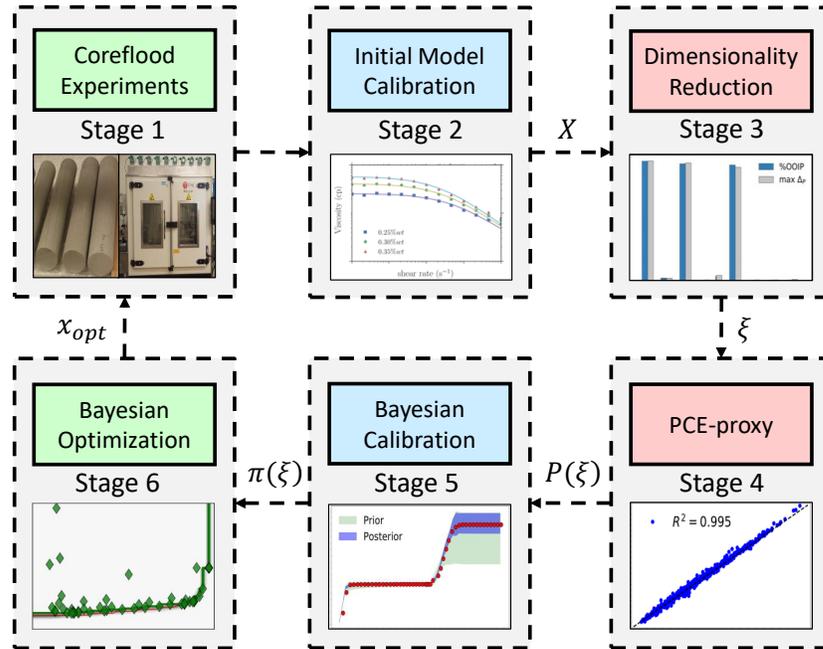
Fig. 4.1. Framework for history matching using Bayesian inference.

Next, we construct a PCE-proxy model in a non-intrusive fashion. This is an important pre-step for Bayesian inference as the posterior distribution of such a high dimensional problem is generally multi-modal non-Gaussian necessitating thousands of evaluations of the expensive simulator [83, 84]. Hence, the prohibitive cost of such an expensive method can be greatly facilitated by using a proxy model that approximates the simulator output response for QoIs. We employ Bayesian inference as it estimates the posterior distribution of the model parameters in a probabilistic sense conditioned on the field data. This is important as this posterior distribution readily quantifies uncertainty in the output response due to the variability in model parameters.

Finally, we pose a multi-objective optimization problem to design an optimized injection strategy for oil recovery using BGO. In BGO, a probabilistic belief is maintained about the objective functions and a surrogate of objective function is constructed using Gaussian processes. This belief is updated using a so-called IAF which determines the next evaluation point for objective functions. In our case, designing this optimized strategy requires consideration and quantification of uncertainty in model parameters which is innate to BGO with extended EIHV IAF [214]. In this work, we show that with only few evaluations of expensive objective functions, implemented BGO is successful in learning the optimal region.

4.4 Uncertainty in SP flooding

The intrinsic variability in the physical properties governing the displacement processes is an inherent characteristic of subsurface flows. This problem is even more pronounced in cEOR due to presence of additional chemo-physical processes. Therefore, successful design of any cEOR strategy depends heavily on how well this uncertainty due to parameter variability is captured and quantified. In this regards, various studies exist in the literature. The effects of variation in surfactant adsorption (ad_{31}) and its implications on the oil recovery have been studied in [50, 177]. The sensitivity of micro-emulsion phase behavior ($C_{33\max 0}, C_{33\max 1}$) on the salinity have been reviewed in [215–217]. Polymer viscosity ($\dot{\gamma}_{1/2}$) and permeability reduction (c_{rk}) was shown to have dominant effect on the sweep efficiency by [132, 218]. The impact of micro-emulsion viscosity (α_{V2}) on oil recovery have been studied by [219, 220]. Here, we quantify uncertainty in these model parameters by considering a $\pm 20\%$ uncertainty bound over the values calibrated from the support experiments which is a reasonable assumption for reservoir problems [132]. For those parameters that were not possible to be inferred from the experiments, we perform an extensive literature review to define a prior range over them (refer to table 4.1).

4.5   Polynomial chaos proxy construction

We adopt the framework from section 3.5 and construct the PCE-proxy in a non-intrusive manner. We use oilcut curve as one of the QoIs in this chapter, as opposed to cumulative oil profile in Chapter 3, since we collect measurement data for the oil-cut curve while the cumulative oil recovery, which can be interpreted as the area under the oilcut curve, is simply calculated from these measured data. This is important in the Bayesian approach as it allows us to capture noise in the measurement process. In this section, we provide the validation of PCE-proxy where the QoIs are oil-cut and pressure profiles.

We first check the accuracy of the PCE-proxy model by plotting *relative L_2 error* in the mean and variance of its output response with the simulation model for our QoIs [195]. The ground truth was established by running simulation model on 1000 LHS experimental design points. Figure 4.2 shows the relative error for pressure drop and oilcut profiles plotted as a function of N_s. Experimental design size is increased till a relative error of 5×10^{-2} is achieved for both the mean and variance plot which was considered to be sufficient for convergence. It is clear from the figures that $N_s = 1000$ is enough to achieve L_2 error within this accuracy for both QoIs.

(a) L₂ error in mean

(b) L₂ error in variance
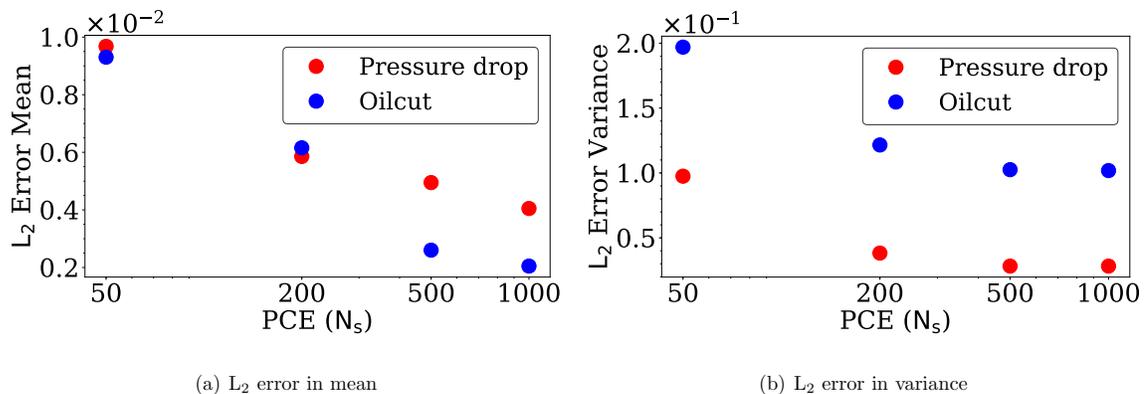
Fig. 4.2. L_2 error in pressure drop and oilcut profiles for PCE-proxy constructed using $N_s = (50, 100, 500, 1000)$ training samples

Figure 4.3 shows the cross-validation of PCE-proxy response with simulation response for the maximum pressure difference and area under the oilcut curve (which can be interpreted as cumulative oil produced) where PCE-proxy is constructed using $N_s = 1000$ experimental design points. It is clearly evident that PCE-proxy with $N_s = 1000$ does an excellent job of approximating this expensive simulation model as the coefficient of determination $R^2 \approx 0.99$ for both of these cases.

(a) $max(\Delta p)$ (b) area under the oilcut curve
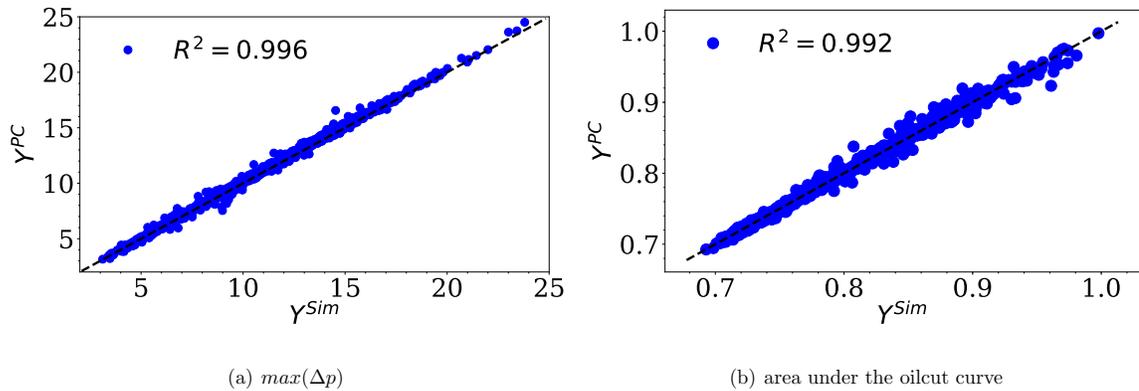
Fig. 4.3. Cross-validation of PCE-proxy (Y^{PC}) response with simulation output (Y^{Sim}) for the maximum pressure difference and area under the oilcut curve

4.6 Bayesian history matching

Bayesian inference is used here for history matching as it provides a framework for a probabilistic model construction for estimation of uncertain model parameters and readily quantifies uncertainty in the model parameters in a probabilistic sense [82]. Bayesian inference requires two main ingredients: 1) a *prior* over model parameters $p(\xi)$ that captures the state of the knowledge about the model parameters before the data is observed. 2) a *likelihood* function $p(D|\xi, M)$ that models the measurement processes; the probability of observing the data D, if the model M assumed is correct. Then as per Bayes' theorem, the *posterior* of the model parameters $p(\xi|D, M)$ condi-

tioned on the observed data can be inferred from the product of prior and likelihood functions: [221, 222]:

$$p(\xi|D, M) = \frac{p(D|\xi, M)p(\xi)}{p(D, M)} \qquad (4.1)$$

This posterior captures the state of the knowledge of the unknown model parameters after the data is observed. We assume a *uniform* priori distribution over all the uncertain model parameters:

$$p(\xi) \sim \mathcal{U}(\prod_{j=1}^{d}[a_j, b_j]) \qquad (4.2)$$

The lower bound a_j and upper bound b_j of the distribution for each parameter ξ_j is provided in table 4.1. We model the measurement process using *Gaussian likelihood* model which can be stated as:

$$p(D|\xi, M) = \mathcal{N}(D_i|f(\xi), \sigma_i^2), \qquad (4.3)$$

where $i = 1, 2$ are QoIs i.e. pressure drop and oilcut profiles. σ_i is the noise parameter associated with measurement process of each QoI. In other words, the likelihood function is constructed such that there is a different noise parameter, σ_i associated with the measurements obtained for each QoI. This noise parameter is modeled with the non-informative Jeffreys' prior which remains invariant under reparametrization [223]. It is straight forward to derive Jeffreys' prior for Gaussian likelihood function as:

$$p(\sigma_i^2) \propto \frac{1}{\sigma_i^2}. \qquad (4.4)$$

$f(\xi)$ in Eq. (4.3) is the forward simulation model (M), which we approximate using PCE-proxy (refer to Eq. (3.4)):

$$f(\xi) \approx P(\xi) \qquad (4.5)$$

$P(D|M)$ in Eq. (4.1), called *normalization constant* is given as:

$$P(D|M) = \int_{\xi} p(D|\xi, M)p(\xi)d\xi \qquad (4.6)$$

The analytical expression for the posterior distribution of such high dimensional problem is not readily available and hence, we resort to the well-known MCMC sampling

technique to sample from the posterior. Here, we use the standard Metropolis-Hashtings (MH) sampler to generate Markov Chains (MC) [224]. In theory, MH generates reversible and ergodic MC that converge to the stationary distribution of uncertain parameters in the limits of infinity [190]. Other algorithms that can be used to generate these MC include Gibbs Sampler [225], Adaptive Metropolis [226], Random Walk Metropolis [227].

We use PyMC package to setup this inverse model calibration problem [228]. We first train the Bayesian model on BCF1 (refer to table 2.5) for our QoIs by generating 10^5 MCMC samples. Appendix B shows the trace of MC samples generated using MH algorithm for all the uncertain parameters. To ensure proper mixing of MC, we *burn* the first 20k samples and then *thin* it out by keeping every 5th sample for estimating the posterior distribution which appears to be sufficient for generating Markov states with low degree of auto-correlation (refer to Appendix C). Appendix D shows the posterior and prior distribution for all the uncertain parameters. An interesting observation is that the posterior for parameters $\dot{\gamma}_{1/2}, C_{33\,\text{max}\,1}$ and ad_{31} looks similar to their priors signifying the importance of initial model calibration stage in the framework. Finally, Geweke [229] posterior diagnostic test ensures that MC has properly convergence for all the parameters (refer to Appendix E).

It can be clearly seen from figure 4.4 that the mean of the posterior does a very job of predicting BCF1 experimental data for all our QoI. The uncertainty observed after PV $\approx$ 1.45 is due to the large uncertain parameter set describing mechanism of SP and P flooding processes. To ensure that Bayesian model has been calibrated accurately, we validate its response for two other experiments BCF2 and BCF3 (refer to table 2.5). From the figure 4.5, prior response space for all QoIs can be clearly seen to be reduced with the estimated posterior accurately predicting the observed data. Next, we use these calibrated models for stochastic optimization under quantified uncertainty in their parameters.

(a) Pressure drop profile (b) Oilcut curve (c) Cumulative oil recovery
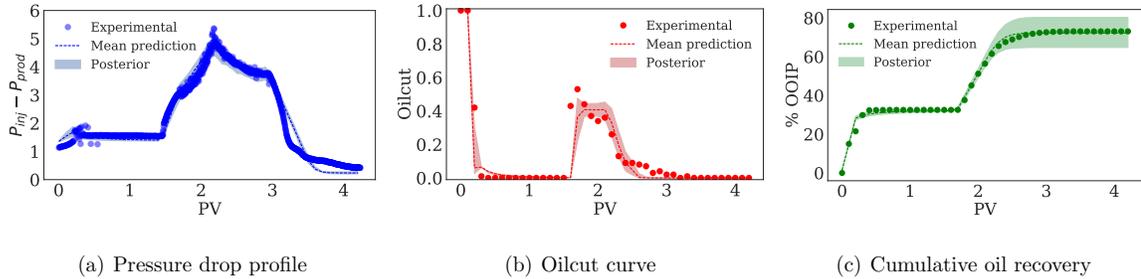
Fig. 4.4. Prediction of output response for BCF1 (training set) using Bayesian inference.

4.7 Stochastic optimization under uncertainty

The injection sequence for oil recovery using SP flooding is very well understood [230] and follows IWF $\rightarrow$ HTDS $\rightarrow$ SP $\rightarrow$ P $\rightarrow$ EWF. However, determining the optimum slug sizes of the injected chemicals is a subject of technical consideration as they govern the physio-chemical processes. Moreover, the injection timing is a subject of economic considerations due to stochastic nature of oil price [231]. To make it even more complicated, the parametric uncertainty which we have quantified in section 4.6 should be taken into account while determining the slug sizes, as it is likely to affect the flow physics. Therefore, we implement BGO to determine the optimum injection slug sizes or timings of chemical slugs under the uncertainties in both the oil prices and model parameters by posing a multi-objective optimization problem. Table 4.2 shows the range of slug sizes (s_{htds}, s_{sp}, s_p) used in this optimization problem where the injection of IWF slug is kept constant at 1.0 PV in all simulations which reduces the residual oil saturation to $S_{or} \approx 0.4$. After this, oil production ceases and cEOR must start.

In reservoir applications, unwarranted pressure changes are unfavorable as they result in stress changes affecting the fracture permeability, and damaging reservoir and wells [123, 124]. Also from the economic perspective, a strategy that produces same oil but requires less pumping power is desirable. Hence, it is important to have

(a) Pressure drop profile

(b) Pressure drop profile

(c) Oilcut curve

(d) Oilcut curve

(e) Cumulative oil recovery

(f) Cumulative oil recovery
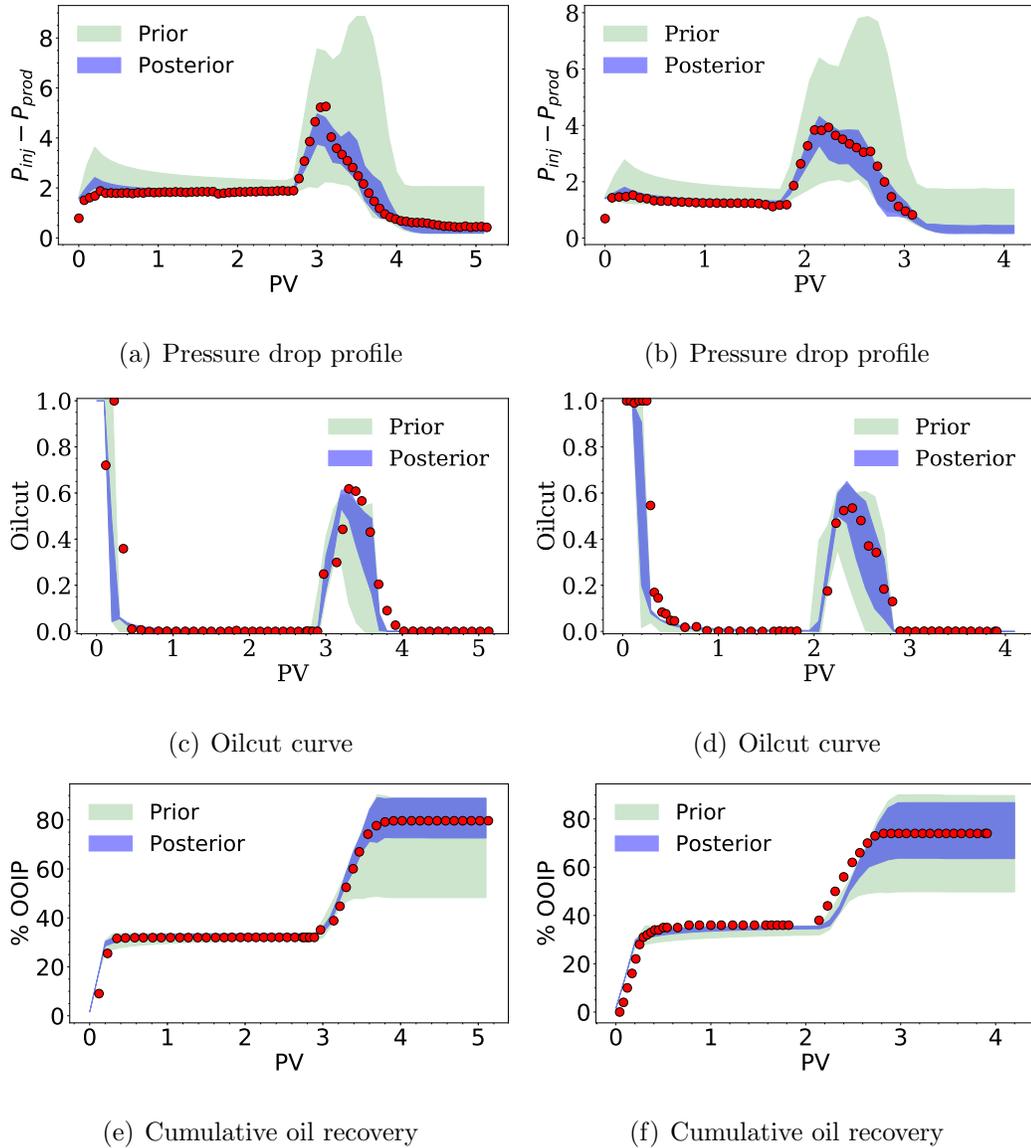
Fig. 4.5. Validation of calibrated Bayesian model with BCF2 and BCF3. The prior and posterior is estimated using 200 simulations

an estimate of pressure change during a cEOR process. Therefore, we use pressure difference between injection and the production well as our fist objective for the optimization problem where the aim is to minimize the maximum pressure difference $max(\Delta p)$.

As stated before, the financial aspect is the most important factor that needs to be considered to gauge success of any cEOR project. Therefore, the second objective we use for the optimization is *Net Present Value* (NPV) which is a direct measure of the profitability [232, 233]. NPV is computed as:

$$NPV = \sum_{j=1}^{N_t} \left[\frac{C_I(\Delta t_j)}{(1+d)^{t_j}} \right] - C_o \tag{4.7}$$

where N_t is the total duration of project, d is the annual discount rate, C_o is the initial capital investment and $C_I(\Delta t_j)$ is the total cash inflow in time Δt_j ($t_{j+1} - t_j = 1$ year). The total cash inflow is calculated as the difference of total revenue or earning and total expenses for the overall duration of the project as [234]:

$$C_I(\Delta t_j) = R(\Delta t_j) - I(\Delta t_j) - T(\Delta t_j) \tag{4.8}$$

$I(\Delta t_j)$ is the total chemical injection cost given as

$$I(\Delta t_j) = c_{salt}(\Delta t_j)S_{salt} + c_{surf}(\Delta t_j)S_{surf} + c_{poly}(\Delta t_j)S_{poly} \tag{4.9}$$

where $c_{salt}, c_{surf}, c_{poly}$ are the mass concentrations of TDS, surfactant and polymer and $S_{salt}, S_{surf}, S_{poly}$ are the costs associated with them respectively. In our optimization, we vary the slug sizes by keeping the total chemical mass concentration constant i.e.

$$c_k = c_k^{ref} \frac{s_k^{ref}}{s_k}, \quad where \; k = surf, \; poly \tag{4.10}$$

where c_k^{ref} is the total mass concentration of injected chemicals at the reference slug size s_k^{ref}. This allows us to design an injection strategy that maximizes the profit with a constant cost of chemical injection.

In Eq. (4.8), $T(\Delta t_j)$ is the total direct expenditure (TDE) which includes expenditure due to operation costs, taxes, water treatment costs and maintenance costs. To simplify the model, $T(\Delta t_j)$ is kept constant as given in table 4.2. Finally, $R(\Delta t_j)$ is the total revenue from oil production:

$$R(\Delta t_j) = S(t_j)N_p(\Delta t_j) \tag{4.11}$$

where $N_p(\Delta t_j)$ is the net barrels of oil produced in Δt_j and $S(t_j)$ is the oil price at time t_j. Oil price is extremely volatile and fluctuations in oil price is natural over long periods of time [234]. Hence, we model oil price using a stochastic oil price model (SOPM) given as $S(t_j) = S(t_o)e^{W_j}$ where W_j is random walk with a drift given by [235]:

$$W_{j+1} = W_j + \alpha_o \xi_{o,j} + \mu. \tag{4.12}$$

where $\xi_{o,j}$ are independent identically distributed standard normal variables. We use initial oil price $S(t_o) = 75$ \$/bbl, $\alpha_o = 10^{-4}$ and $\mu = \pm 10^{-5}$. μ models the drift in the oil price. Here, we investigate the influence of both positive and negative drifts in the oil price on the optimized injection sequence (refer to figure 4.6). Table 4.2 shows the details of the optimization problem and the reference values used in the NPV model.

(a) $\mu = +10^{-5}$

(b) $\mu = -10^{-5}$
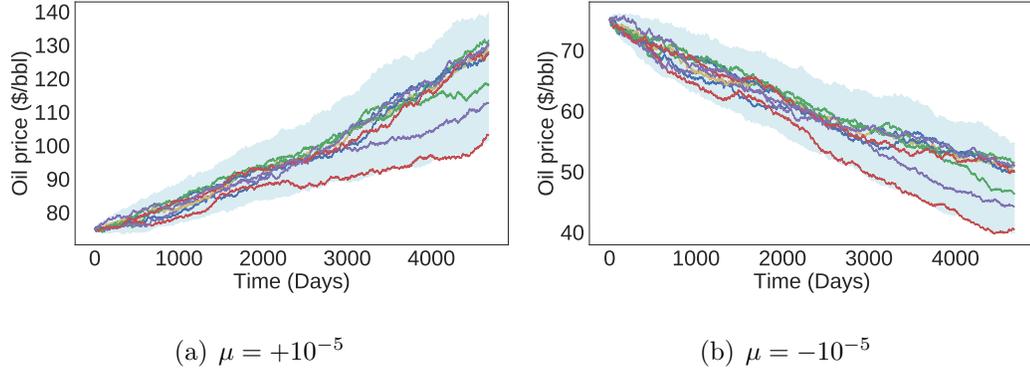
Fig. 4.6. Random samples from the stochastic oil price model for a) positive drift, b) negative drift

A statistical design methodology [214] is applied to solve this two-objective stochastic optimization problem. We formulate the problem as follows:

$$\hat{\mathbf{x}} = \arg\left(\max_{\mathbf{x}} \mathbb{E}[O_1(\mathbf{x}; \xi)], \min_{\mathbf{x}} \mathbb{E}[O_2(\mathbf{x}; \xi)]\right), \tag{4.13}$$

where $\mathbf{x}$ is the vector of design variables i.e. the slug sizes. We maximize the expectation $\mathbb{E}[\cdot]$ over the noise for the first objective (i.e. $O_1(\mathbf{x})$ is NPV) and minimize

Table 4.2.
Design of optimization problem. Note that the salt concentration is kept constant irrespective of the slug size owing to its low cost.

	Quantity	Variable	Value	Range	Units
HTDS Slug	Salt conc.	c_{salt}	17500	const.	ppm
	Slug size	s_{htds}	-	0.1-1.0	PV
SP Slug	Salt conc.	c_{salt}	17500	const.	ppm
	Surf conc.	c_{surf}^{ref}	0.68	Eq. (4.10)	%
	Poly conc.	c_{poly}^{ref}	25000	Eq. (4.10)	ppm
	Slug size.	s_{surf}^{ref}	0.25	0.1-1.0	PV
P Slug	Salt conc.	c_{salt}	9400	const.	ppm
	Poly conc.	c_{poly}^{ref}	25000	Eq. (4.10)	ppm
	Slug size.	s_{poly}^{ref}	0.25	0.1-1.0	PV
Cost	Capital	C_o	200000	-	$
	Salt price	S_{salt}	0.5	-	$/lb
	Surf price	S_{surf}	5	-	$/lb
	Poly price	S_{poly}	2	-	$/lb
	TDE	$T(\Delta t_j)$	10000	-	$/year
Misc.	Discount rate	d	20	-	%
	Total time	N_t	13	-	years

for the second (i.e. $O_2(\mathbf{x})$ is $max(\Delta p)$). Due to the lack of an *apriori* preference structure, the two objectives may be competing, i.e. choosing a process input that increases NPV may result in an increased value of $max(\Delta p)$. We say that a set of designs $\mathbf{x}_1$ dominates a different set of designs $\mathbf{x}_2$ if $\mathbf{x}_1$ results in a higher NPV and a lower $max(\Delta p)$. For any dominated point on the PF, there is always another point on the PF that results in a higher value of NPV and a lower value of $max(\Delta p)$. Since, we maximize NPV and minimize $max(\Delta p)$, we aim to explore points in the

bottom-right region of the hypervolume. The demonstration of the methodology can be seen through the evolution of the PF and its uncertainty in Fig. 4.7 and Fig. 4.8. For both the cases, positive and negative drifts, we start the methodology with 20 initial data points selected using LHS [236].

The state of the PF can be seen after the first, 50th and the 100th iteration for positive and negative drifts in Fig. 4.7 and Fig. 4.8, respectively. The green staircase-like line joining the diamonds that are Pareto-optimal represents the PF. The red dot is the point selected by the algorithm at the given stage of the algorithm. The uncertainty around the PF is represented by the shaded grey regions around the PF. The intensity of the region is representative of how plausible it is to discover a Pareto-optimal point in that region, as quantified by the colorbar accompanying Fig. 4.7 and Fig. 4.8. It is intuitive that, as more BGO guided simulations points are added, the shaded grey areas around the PF become thinner and collapse on the PF. This means that the algorithm believes that there exists little chance of discovering more Pareto-optimal points in regions further towards the bottom-right part of the hypervolume. This visual inspection is one criterion that can be used to stop sampling. More generally, a predetermined number of simulations is used as a stopping criterion. For more details of the methodology, we refer interested readers to [214].

(a) $N_{itr} = 1$ (b) $N_{itr} = 50$ (c) $N_{itr} = 100$
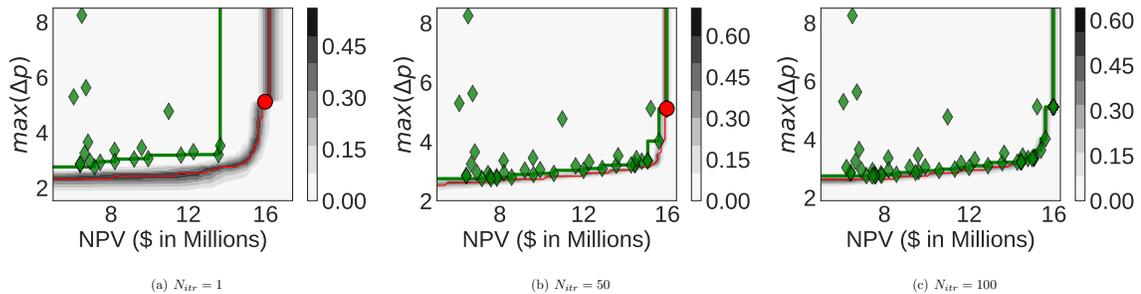
Fig. 4.7. Pareto front for positive drift stochastic oil price model after iterations $N_{itr} = 1, 50, 100$
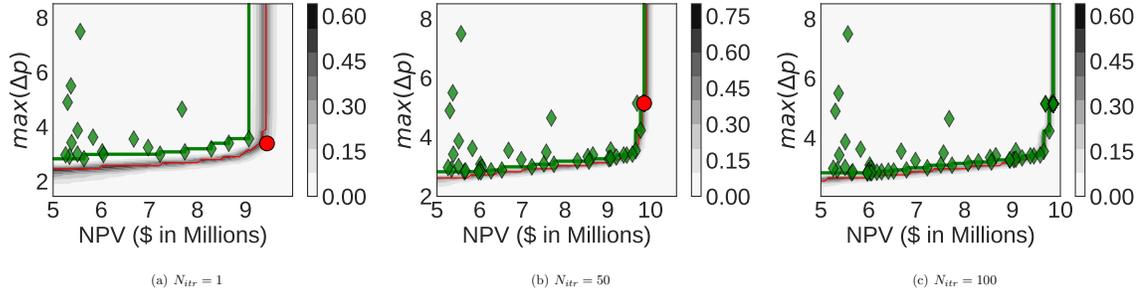
Fig. 4.8. Pareto front for negative drift stochastic oil price model after iterations $N_{itr} = 1, 50, 100$

Figures 4.9-4.10 shows the slug size design with parametric uncertainty for the selected point (called as knee-design point) from PF using minimum distance selection method (TMDSM) [237,238]. It is interesting to see that the algorithm quickly learns that for the positive drift case, injection of chemicals should be as delayed as possible. This is because as the oil price is always increasing for this case, the revenue generated from oil production (refer to Eq. (4.11)) would be more if the same oil is produced late ($s_{htds} = 1.0$PV, $s_{surf} = 0.3$PV, $s_{poly} = 0.4$PV). Whereas, for the negative drift case where oil price is always diminishing, it is advisable to inject chemicals as early as possible when the oil price is still high to maximize the profit ($s_{htds} = 0.1$PV, $s_{surf} = 0.24$PV, $s_{poly} = 0.88$PV). We use 50 MC samples from the estimated posterior in section 4.6 to quantify parametric uncertainty in figures 4.9-4.10.
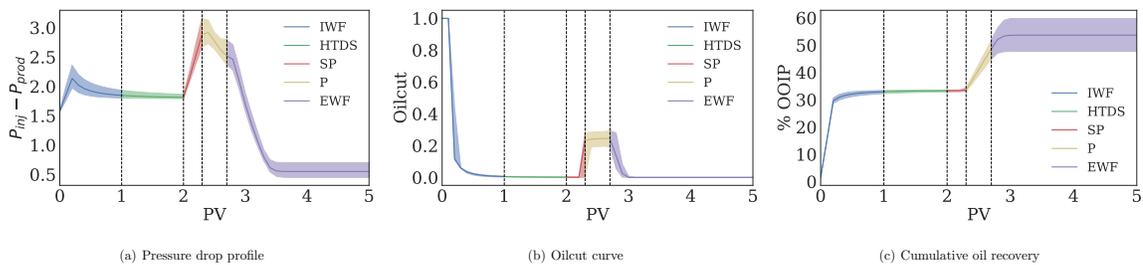

Fig. 4.9. Uncertainty quantification for the knee-design point selected from the Pareto front of positive drift SOPM having the slug design as: $s_{htds} = 1.0$PV, $s_{surf} = 0.3$PV, $s_{poly} = 0.4$PV.

(a) Pressure drop profile   (b) Oilcut curve   (c) Cumulative oil recovery
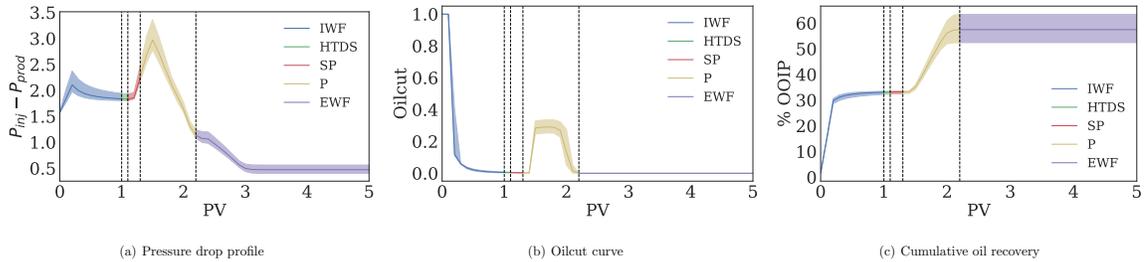
Fig. 4.10. Uncertainty quantification for the knee-design point selected from the Pareto front of negative drift SOPM having the slug design as: $s_{htds} = 0.1$PV, $s_{surf} = 0.24$PV, $s_{poly} = 0.88$PV.

4.8 Conclusions

In this chapter, we presented a Bayesian history matching framework for model calibration of elegant yet complex mechanism of oil recovery through SP flooding. At first, an effective parametrization of prior model parameters is inferred with the help of support-experiments. The resulting high dimensional input parameter space is then reduced with Sobol SA [195]. The forward simulation model response is approximated using PCE-proxy which replaces the expensive physics-based model in likelihood function. MCMC sampling technique is adapted to sample from the resulting non-Gaussian multi-modal posterior distribution. This framework is first implemented to infer the posterior by training the Bayesian model on a single core-flood experiment. The resulting posterior is then tested and validated by accurately predicting QoIs for two other coreflood experiments.

The framework is then extended to a multi-objective optimization problem of injection slug design under uncertainty in the calibrated model parameters with the objectives of minimizing maximum pressure drop and maximizing NPV. A stochastic model for oil price is considered for two cases of positive and negative drifts in the oil price. BGO is adapted to solve this multi-objective optimization problem which builds a probabilistic surrogate of objectives to update the IAF using Gaus-

sian processes. We use EEIHV as acquisition function as it filters out the parametric uncertainties thus bypassing the need to learn surface response of high-dimensional uncertain parameter space. We show that BGO successfully optimizes our objectives by quantifying uncertainty about the predicted Pareto front.

A next step would be to apply this methodology for a more realistic and computationally demanding heterogeneous reservoir problem with an uncertainty in spatial permeability or porosity field such as [193, 239, 240] given the fact that BGO with EEIHV IAF has been solely developed for expensive function optimization. To summarize, the methodology presented could be extended to optimize or calibrate models for more complex history matching problems for variety of applications such as fractured reservoirs [171, 172], CO_2-EOR and storage [174, 175] or alkaline-surfactant-polymer (ASP) flooding [173].

4.9   Computer Code Availability

UQLab is an open source uncertainty quantification framework developed at ETH Zurich (Switzerland) and can be freely downloaded with example data and user manual from `http://www.uqlab.com/download`. The code for BGO with EEIHV IAF is available as free software at `https://github.com/piyushpandita92/pydes`.

4.10   Acknowledgment

This research is partially supported by a grant from the Pioneer Oil Company. The authors would like to thank Bryan Clayton for his support and useful discussion.

5. SUMMARY

Surfactant-polymer (SP) flooding is an important enhanced oil recovery technique which aims at improving mobilization of oil by injecting surfactant and polymer slugs in oil reservoirs. In SP flood, surfactant reduces the interfacial tension between oleic and aqueous phases while polymer enhances the mobility control by inhibiting the viscous fingering instabilities thereby improving oil recovery. The underlying physical models governing each of these sub-processes can be calibrated by performing lab-scale coreflood experiments. The performance of SP flood on field-scale is conditioned on predictive capabilities of these models which depends on how well these models are calibrated. Moreover, model predictions are usually subject to epistemic uncertainty in experimental measurements or parametric uncertainty in the model calibration stage. Therefore, these uncertainties in high-dimensional space of model parameters should be addressed in the model calibration stage.

In this regard, this thesis presents a comprehensive study on uncertainty quantification and history matching of surfactant-polymer flooding. At first, we construct a mechanistic SP flood simulation model by performing extensive lab-scale experiments. We then introduce a framework for model calibration which involves sequential execution of sensitivity analysis, proxy modeling and inverse optimization to determine the optimized parameter space of model parameters. The employed model calibration algorithm starts with Sobol sensitivity analysis which reduces the large uncertain parameter space of model parameters to find the most important stochastic variables that greatly affects the output response of simulation model. This low-dimensional parameter space is then used to construct an efficient proxy model based on polynomial chaos expansion (PCE-proxy). PCE-proxy approximates the highly non-linear output response of simulation model and guarantees convergence in distribution of output quantities of interest. The error in PCE-proxy posterior is estimated using

empirical norm error and its variant leave-one-out error which is shown to converge with an increase in the number of training data points. The moments of PCE-proxy i.e. mean and variance is compared with the moments of stochastic approach with 10^4 LHS simulations. It is shown that even for the PCE-proxy constructed with 200 (N_s) data points, the relative L^2 error in the moments converge to an acceptable value.

We then posed a multi-objective inverse optimization problem using PCE-proxy to history match experimental observations. This problem is solved using the Genetic algorithm (GA) to obtain an optimized subspace of model parameters that minimizes the miss-fit between PCE-proxy model response and observations. We use this framework to first calibrate a single coreflood experiment for quantities of interest such as pressure profile, cumulative oil recovery curve and oilcut curve. We then show that the calibrated model is able to predict all our quantities of interest (QoIs) for two other coreflood experiments without any further tuning of these model parameters. Finally, we combine PCE-proxy with Gaussian regression process or Kriging to quantify epistemic uncertainty in the proxy model.

Next, we extend this framework for Bayesian history matching where the probability distribution of the uncertain model parameters is inferred by directly sampling from their posterior using MCMC, the acceptance criteria of which is tested using the standard Metropolis-Hasting (MH) algorithm. To effectively sample using MCMC, we replace the expensive physics based forward model in the Gaussian likelihood function with the cheap PCE-proxy. We use a single coreflood experiment to calibrate the Bayesian model and validate its convergence by successfully predicting the QoIs for two other coreflood experiments without any ad-hoc tuning of the posterior distribution. We then perform a multi-objective optimization of injection slug size design under uncertainties in model parameters and oil price. We implement Bayesian global optimization (BGO) for this stochastic multi-objective optimization problem which is a class of algorithms that has found immense use in designing computer simulations that are expensive while remaining non-intrusive and gradient free. BGO models the objectives via Gaussian process surrogates and uses the epistemic uncertainty to

define an information acquisition function that quantifies the merit of evaluating the objective at new design points. We employ the extended expected improvement over the dominated hypervolume (EEIHV) as an information acquisition function as it readily quantifies noise in the model parameters, bypassing the need of learning their response on the objectives. We adapt random walk with a drift to model stochastic oil price for two cases of positive and negative drift and show that the implemented BGO successfully exploits the optimum regions of the objectives for both of these cases characterizing confidence about the predicted Pareto front. It is concluded that for the positive drift, when the oil price is always increasing, the algorithm predicts a delayed slug size scenario to maximize the profit. However, for the case of negative drift, when the oil price is always decreasing, the algorithm predicts smaller slug size scenario to minimize the loss. The presented proxy-accelerated history matching and stochastic optimization approach could be extended to more complex problems outside the realm of reservoir engineering.

APPENDICES

A. PSEUDO-CODE

This appendix provides the psuedo-code for sensitivity analysis, PCE-proxy modeling and model calibration using Genetic Algorithm.

Algorithm 1 Pseudo code for sensitivity analysis

1: **procedure** SENSITIVITY ANALYSIS FOR DIMENSION REDUCTION

2:     define models for each sub-process $T_{models} = (M_1, M_2...M_t)$

3: *loop*:

4:     initial model calibration using experiments or literature (M_i)

5:     consider $\pm$ 20% uncertainty and uniform distribution on input parameters $X_{Mi} = (X_{1Mi}, X_{2Mi}...X_{nMi})^T$ for M_i

6:     generate $N_{samples}$ of X_{Mi}

7:     perform Sobol sensitivity analysis

8:     get the reduced input vector (here $m \leq n$) $\xi_{Mi} = (\xi_{1Mi}, \xi_{2Mi}...\xi_{mMi})^T$

9:     **goto** *loop* until M_t

10:     Combine all $\xi = (\xi_{M1}, \xi_{M2}...\xi_{Mn})^T$

11: **end procedure**

Algorithm 2 Pseudo code for PCE-proxy modeling

    procedure PCE PROXY MODELING

2:      **if** $\Delta\epsilon_{gen}^{mean} > \epsilon_{accept}$ **then** increase training points (N_s)

        define input design matrix $\xi = (\xi_1, \xi_2...\xi_n)^T$

4:      evaluate computational model $Y = (P_{\xi 1}, P_{\xi 2}...P_{\xi n})^T$

        construct PCE-proxy

6:        **for** $p = p_{min} : p_{max}$ $(degree)$

          solve least square problem

8:        calculate $\Delta\epsilon_{gen}^{mean}$

       end if computational cost reached

10:     get PCE-proxy $P(\xi)$

    end procedure

Algorithm 3 Pseudo code for model calibration

    procedure GENETIC ALGORITHM FOR MODEL CALIBRATION

2:      $N = 0$, $N_{cross} = 0$,

      specify N_{gen}^{max}

4:      **if** $\sigma_{loss} >$ error **then** increase generations (N_{gen}^{max})

        specify pop_{size}

6:        evaluate PCE-proxy $P(\xi)$ at pop_{size}

        $N_{cross} = N_{cross} + n_{cross}$

8:        **while** N $< N_{cross}$

          calculate loss function and select : $S(N) \in P(N)$

10:        generate offspring : $O(N) \leftarrow$ by crossover, mutation$(S(N))$

          new population : $P(N + 1) = \text{combine}(P(N), O(N))$

12:        $N = N + 1$

          get Pareto front (ξ_{min})

14:        evaluate BCF2 and BCF3 at ξ_{min} using simulation model

          calculate σ_{loss}

16:      **end if** N_{gen}^{end} reached

      get ξ_{min} as final solution

18: **end procedure**

B. TRACE OF MCMC SAMPLES
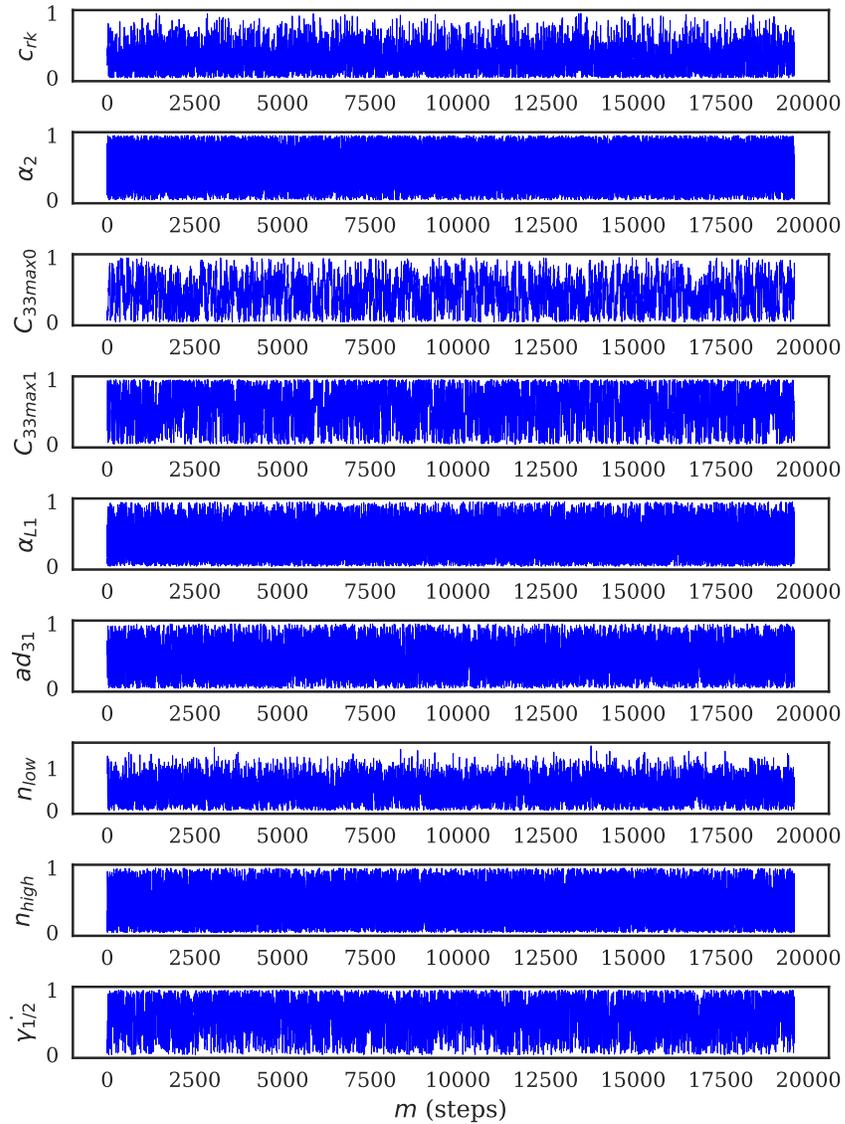
Fig. B.1. Trace of MCMC samples generated using MH algorithm for all the model calibration parameters given in table 4.1

C. AUTO-CORRELATION PLOTS

The degree of auto-correlation can be calculated using the following equation:

$$\rho_k = \frac{Cov(X_t, X_{t+k})}{\sqrt{\mathbb{V}(X_t)\mathbb{V}(X_{t+k})}} \tag{C.1}$$

where $Cov(X_t, X_{t+k})$ is the covariance between the traces of uncertain variables separated by k steps.
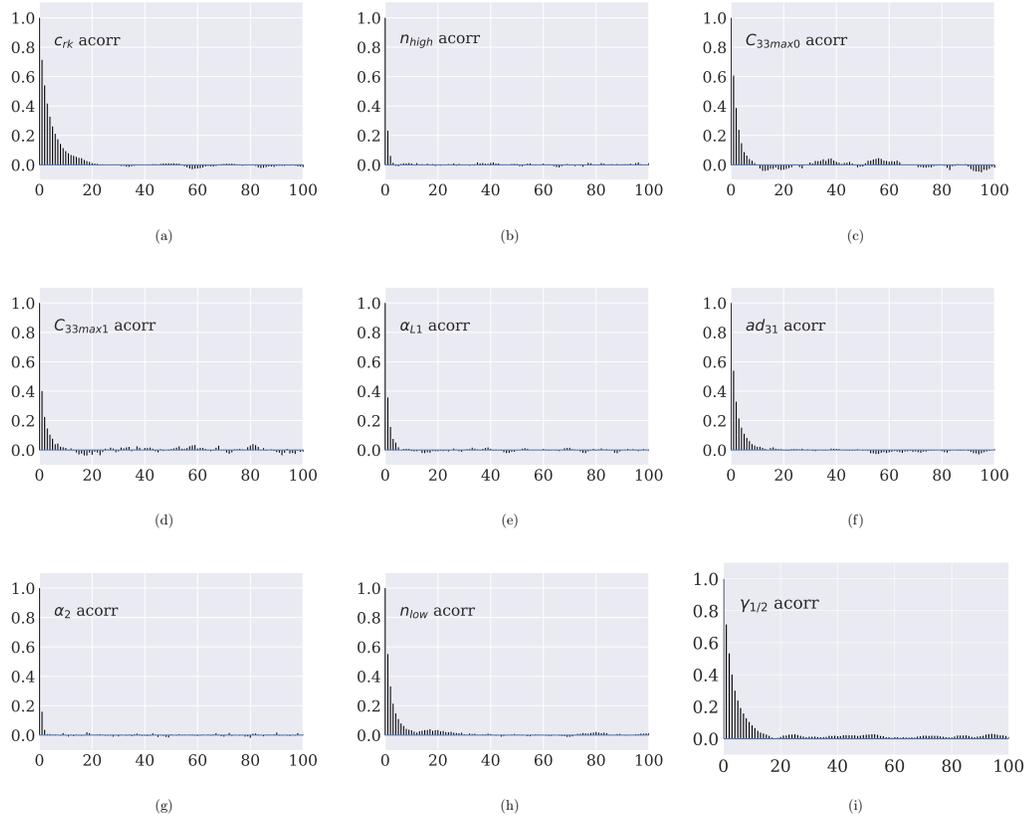
Fig. C.1. Auto-correlation plots for all the uncertain model parameters

D. POSTERIOR DISTRIBUTION

Figure D.1 shows the prior and posterior distribution of all the model parameters.
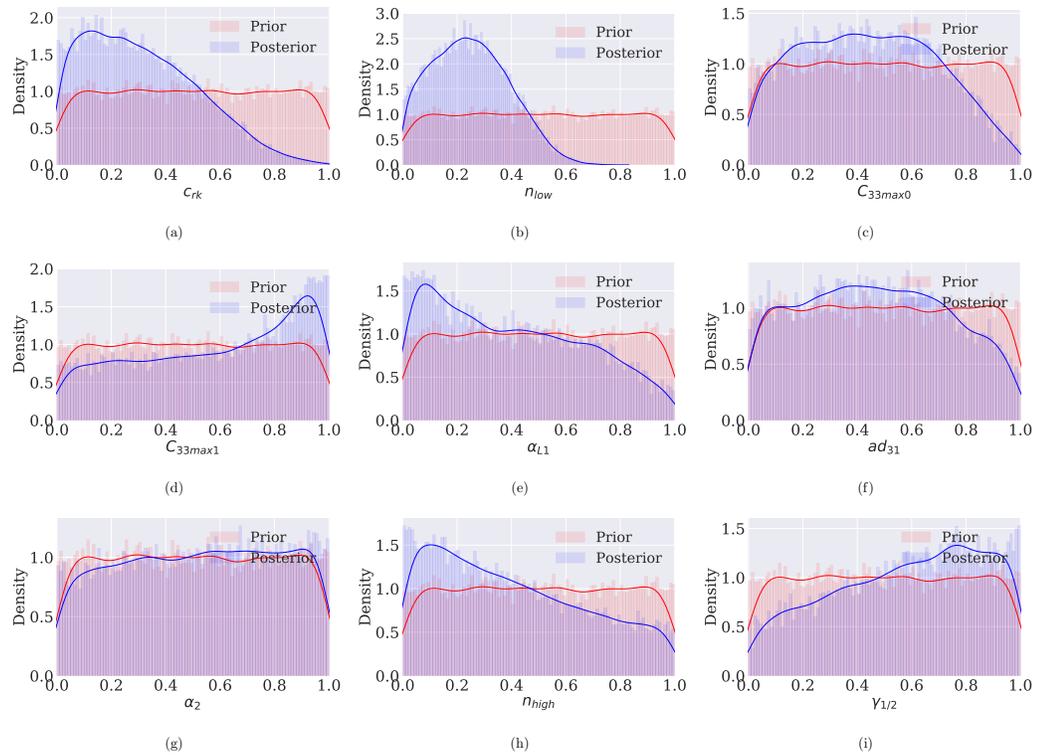
Fig. D.1. Prior and posterior distribution of all the uncertain model parameters.

E. POSTERIOR DIAGNOSTIC

We compare the mean $(\hat{\xi})$ and variance $(\mathbb{V}(\xi))$ of segments from the beginning and end of a single chain as proposed by Geweke [229] and compute the *z-score* as:

$$z = \frac{\hat{\xi}_a - \hat{\xi}_b}{\sqrt{\mathbb{V}(\xi_a) + \mathbb{V}(\xi_b)}} \tag{E.1}$$

where $a = 10\%$ of initial chain length and $b = 50\%$ of remaining chain. From figure E.1 it is evident that MC has converged for all uncertain variables as the majority of points (computed for various sub-chains) are within 2 standard deviations of zero.
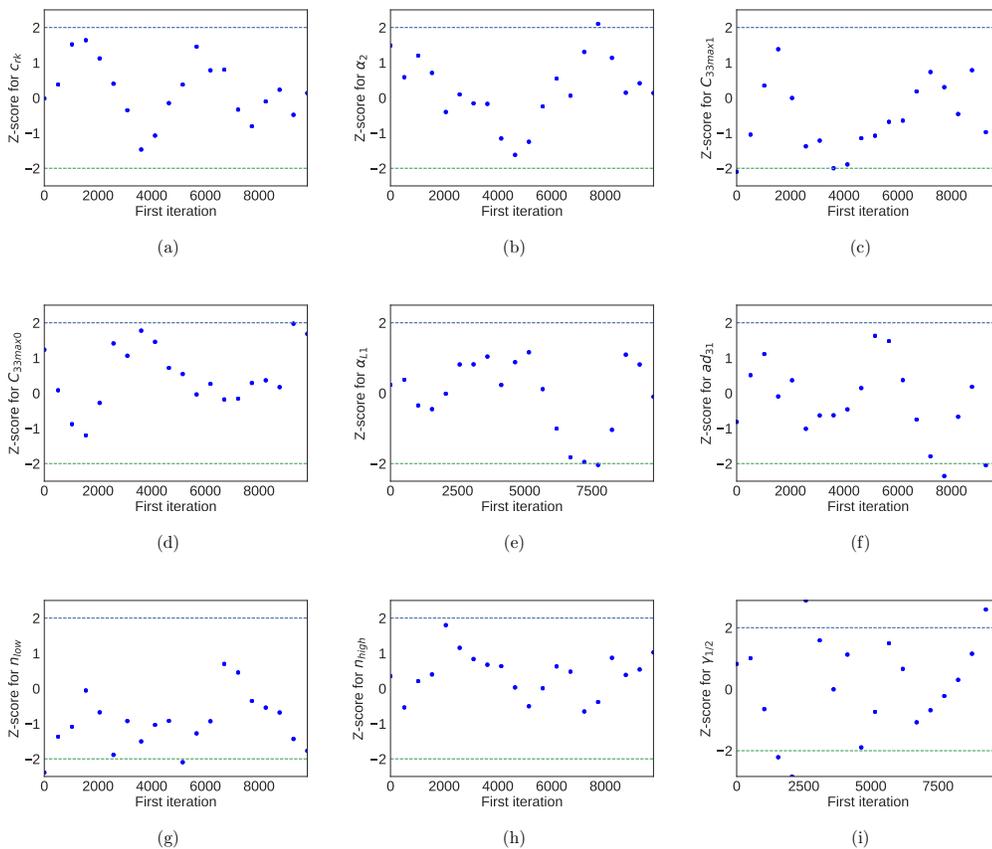
Fig. E.1. Geweke-plot for posterior diagnostic of all the model parameters.

REFERENCES

REFERENCES

[1] A. Gurgel, M. Moura, T. Dantas, E. B. Neto, and A. D. Neto, "A review on chemical flooding methods applied in enhanced oil recovery," *Brazilian journal of petroleum and gas*, vol. 2, no. 2, 2008.

[2] C. Negin, S. Ali, and Q. Xie, "Most common surfactants employed in chemical enhanced oil recovery," *Petroleum*, vol. 3, no. 2, pp. 197–211, 2017.

[3] J. J. Sheng, "Surfactant enhanced oil recovery in carbonate reservoirs," in *Enhanced Oil Recovery Field Case Studies*. Elsevier, 2013, pp. 281–299.

[4] J. Sheng, *Modern chemical enhanced oil recovery: theory and practice*. Gulf Professional, 2016.

[5] L. W. Lake, *Fundamentals of enhanced oil recovery*. Society of Petroleum Engineers, 2014.

[6] D. W. Green, G. P. Willhite *et al.*, *Enhanced oil recovery*. Henry L. Doherty Memorial Fund of AIME, Society of Petroleum Engineers Richardson, TX, 1998, vol. 6.

[7] V. Alvarado and E. Manrique, "Enhanced oil recovery: an update review," *Energies*, vol. 3, no. 9, pp. 1529–1575, 2010.

[8] A. A. Olajire, "Review of asp eor (alkaline surfactant polymer enhanced oil recovery) technology in the petroleum industry: Prospects and challenges," *Energy*, vol. 77, pp. 963–982, 2014.

[9] M. S. Kamal, I. A. Hussein, and A. S. Sultan, "Review on surfactant flooding: phase behavior, retention, ift, and field applications," *Energy & Fuels*, vol. 31, no. 8, pp. 7701–7720, 2017.

[10] S. B. Sandersen, E. H. Stenby, and N. von Solms, "Enhanced oil recovery with surfactant flooding," 2012.

[11] M. Bourrel and R. S. Schechter, *Microemulsions and related systems: formulation, solvency, and physical properties*. Editions Technip, 2010.

[12] M. Bhattarai, "A numerical modeling study of surfactant enhanced mobilization of residual lnapl using utchem." 2006.

[13] J. J. Sheng, "Chapter 5 - surfactantpolymer flooding," in *Enhanced Oil Recovery Field Case Studies*, J. J. Sheng, Ed. Boston: Gulf Professional Publishing, 2013, pp. 117 – 142. [Online]. Available: http://www.sciencedirect.com/science/article/pii/B9780123865458000051

[14] A. Kogan and N. Garti, "Microemulsions as transdermal drug delivery vehicles," *Advances in colloid and interface science*, vol. 123, pp. 369–385, 2006.

[15] K. Raney and C. Miller, "Diffusion path analysis of dynamic behavior of oil-water-surfactant systems," *AIChE journal*, vol. 33, no. 11, pp. 1791–1799, 1987.

[16] M. Mohammed and T. Babadagli, "Wettability alteration: A comprehensive review of materials/methods and testing the selected ones on heavy-oil containing oil-wet systems," *Advances in colloid and interface science*, vol. 220, pp. 54–77, 2015.

[17] J. J. Sheng, "Review of surfactant enhanced oil recovery in carbonate reservoirs," *Advances in Petroleum Exploration and Development*, vol. 6, no. 1, pp. 1–10, 2013.

[18] ——, "Comparison of the effects of wettability alteration and ift reduction on oil recovery in carbonate reservoirs," *Asia-Pacific Journal of Chemical Engineering*, vol. 8, no. 1, pp. 154–161, 2013.

[19] G. Stegemeier, "Mechanisms of entrapment and mobilization of oil in porous media," in *Improved oil recovery by surfactant and polymer flooding*. Elsevier, 1977, pp. 55–91.

[20] B. Song, X. Hu, X. Shui, Z. Cui, and Z. Wang, "A new type of renewable surfactants for enhanced oil recovery: dialkylpolyoxyethylene ether methyl carboxyl betaines," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 489, pp. 433–440, 2016.

[21] S. Kumar and A. Mandal, "Studies on interfacial behavior and wettability change phenomena by ionic and nonionic surfactants in presence of alkalis and salt for enhanced oil recovery," *Applied Surface Science*, vol. 372, pp. 42–51, 2016.

[22] A. M. Howe, A. Clarke, J. Mitchell, J. Staniland, L. Hawkes, and C. Whalan, "Visualising surfactant enhanced oil recovery," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 480, pp. 449–461, 2015.

[23] J. Sukpisan, J. Kanatharana, A. Sirivat, and S. Wang, "The specific viscosity of partially hydrolyzed polyacrylamide solutions: Effects of degree of hydrolysis, molecular weight, solvent quality and temperature," *Journal of Polymer Science Part B: Polymer Physics*, vol. 36, no. 5, pp. 743–753, 1998.

[24] D. Wever, F. Picchioni, and A. Broekhuis, "Polymers for enhanced oil recovery: a paradigm for structure–property relationship in aqueous solution," *Progress in Polymer Science*, vol. 36, no. 11, pp. 1558–1628, 2011.

[25] A. Abidin, T. Puspasari, and W. Nugroho, "Polymers for enhanced oil recovery technology," *Procedia Chemistry*, vol. 4, pp. 11–16, 2012.

[26] G. Jerauld, H. Davis, L. Scriven *et al.*, "Stability fronts of permanent form in immiscible displacement," in *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 1984.

[27] K. S. M. Delshad, G. Pope, "Utchem version 9.82 technical documentation," *Center for Petroleum and Geosystems Engineering*, 2000.

[28] S. Aramideh, P. P. Vlachos, and A. M. Ardekani, "Unstable displacement of non-aqueous phase liquids with surfactant and polymer," *Transport in Porous Media*, pp. 1–20, 2018.

[29] J. Bear, "Dynamics of fluids in porous materials," *Society of Petroleum Engineers: Dallas, TX, USA*, 1972.

[30] D. W. Peaceman *et al.*, "Interpretation of well-block pressures in numerical reservoir simulation with nonsquare grid blocks and anisotropic permeability," *Society of Petroleum Engineers Journal*, vol. 23, no. 03, pp. 531–543, 1983.

[31] D. Babu, A. Odeh, A. Al-Khalita, R. McCann *et al.*, "Supplement to spe 20161, the relation between wellblock and wellbore pressure in numerical simulation of horizontal wells–general formulas for arbitrary well locations in grids," 1991.

[32] J. Kou and S. Sun, "On iterative impes formulation for two phase flow with capillarity in heterogeneous porous media," *International Journal of Numerical Analysis and Modeling. Series B*, vol. 1, no. 1, pp. 20–40, 2010.

[33] J. Franc, P. Horgue, R. Guibert, and G. Debenest, "Benchmark of different cfl conditions for impes," *Comptes Rendus Mécanique*, vol. 344, no. 10, pp. 715–724, 2016.

[34] P. Horgue, C. Soulaine, J. Franc, R. Guibert, and G. Debenest, "An open-source toolbox for multiphase flow in porous media," *Computer Physics Communications*, vol. 187, pp. 217–226, 2015.

[35] M. Karimi-Fard, A. Firoozabadi *et al.*, "Numerical simulation of water injection in 2d fractured media using discrete-fracture model," in *SPE annual technical conference and exhibition*. Society of Petroleum Engineers, 2001.

[36] J. Monteagudo and A. Firoozabadi, "Control-volume method for numerical simulation of two-phase immiscible flow in two-and three-dimensional discrete-fractured media," *Water resources research*, vol. 40, no. 7, 2004.

[37] A. Harten, "High resolution schemes for hyperbolic conservation laws," *Journal of computational physics*, vol. 49, no. 3, pp. 357–393, 1983.

[38] S. E. Buckley, M. Leverett *et al.*, "Mechanism of fluid displacement in sands," *Transactions of the AIME*, vol. 146, no. 01, pp. 107–116, 1942.

[39] H. J. Welge *et al.*, "A simplified method for computing oil recovery by gas or water drive," *Journal of Petroleum Technology*, vol. 4, no. 04, pp. 91–98, 1952.

[40] Y. Al-Wahaibi, C. Grattoni, and A. Muggeridge, "Drainage and imbibition relative permeabilities at near miscible conditions," *Journal of Petroleum Science and Engineering*, vol. 53, no. 3-4, pp. 239–253, 2006.

[41] A. Riaz and H. A. Tchelepi, "Influence of relative permeability on the stability characteristics of immiscible flow in porous media," *Transport in porous media*, vol. 64, no. 3, pp. 315–338, 2006.

[42] G.-Q. Tang and A. Kovscek, "High resolution imaging of unstable, forced imbibition in berea sandstone," *Transport in porous media*, vol. 86, no. 2, pp. 617–634, 2011.

[43] A. Riaz and H. A. Tchelepi, "Numerical simulation of immiscible two-phase flow in porous media," *Physics of Fluids*, vol. 18, no. 1, p. 014104, 2006.

[44] Y. Yortsos and F. Hickernell, "Linear stability of immiscible displacement in porous media," *SIAM Journal on Applied Mathematics*, vol. 49, no. 3, pp. 730–748, 1989.

[45] J. J. Sheng, "Surfactant-polymer flooding," *Modern Chemical Enhanced Oil Recovery*, pp. 371–387, 2011.

[46] D. O. Shah and R. S. Schechter, *Improved oil recovery by surfactant and polymer flooding.* Academic Press, 1977.

[47] A. Abidin, T. Puspasari, and W. Nugroho, "Polymers for enhanced oil recovery technology," *Procedia Chemistry*, vol. 4, p. 1116, 2012.

[48] M. S. Kamal, I. A. Hussein, and A. S. Sultan, "Review on surfactant flooding: Phase behavior, retention, ift, and field applications," *Energy & Fuels*, vol. 31, no. 8, pp. 7701–7720, 2017. [Online]. Available: https://doi.org/10.1021/acs.energyfuels.7b00353

[49] J. J. Sheng, "Status of surfactant eor technology," *Petroleum*, vol. 1, no. 2, pp. 97–105, 2015.

[50] C. Brown and P. Smith, "The evaluation of uncertainty in surfactant eor performance prediction," *Proceedings of SPE Annual Technical Conference and Exhibition*, 1984.

[51] S. Thomas, "Chemical eor: The past - does it have a future?" *Society of Petroleum Engineers*, vol. 108828, January 2006.

[52] A. Alkhatib and M. Babaei, "Applying the multilevel monte carlo method for heterogeneity-induced uncertainty quantification of surfactant/polymer flooding," *SPE Journal*, vol. 21, no. 04, pp. 1192–1203, Jan 2016.

[53] A. M. Al-Sofi, J. S. Liu, and M. Han, "Numerical simulation of surfactant-polymer coreflooding experiments for carbonates," *SPE EOR Conference at Oil and Gas West Asia*, 2012.

[54] W. Li, D. Zhang, and G. Lin, "A surrogate-based adaptive sampling approach for history matching and uncertainty quantification," *SPE Reservoir Simulation Symposium*, 2015.

[55] F. T. Wall, "Principles of polymer chemistry. paul j. flory. cornell univ. press, ithaca, new york, 1953. 688 pp. illus," *Science*, vol. 119, no. 3095, pp. 555–556, 1954.

[56] M. Ruecker, W.-B. Bartels, E. Unsal, S. Berg, N. Brussee, A. Coorn, and A. Bonnin, "The formation of microemulsion at flow conditions in rock," 08 2017.

[57] R. Nelson and G. Pope, "Phase relationships in chemical flooding," *Society of Petroleum Engineers Journal*, vol. 18, no. 05, p. 325338, Jan 1978.

[58] L. Prouvost, T. Satoh, K. Sepehrnoori, and G. Pope, "A new micellar phase-behavior model for simulating systems with up to three amphiphilic species," *SPE Annual Technical Conference and Exhibition*, 1984.

[59] V. Joekar-Niasar and S. M. Hassanizadeh, "Uniqueness of specific interfacial areacapillary pressuresaturation relationship under non-equilibrium conditions in two-phase porous media flow," *Transport in Porous Media*, vol. 94, no. 2, p. 465486, 2012.

[60] I. Chatzis and N. R. Morrow, "Correlation of capillary number relationships for sandstone," *Society of Petroleum Engineers Journal*, vol. 24, no. 05, p. 555562, Jan 1984.

[61] N. Walton, "Electrical conductivity and total dissolved solidswhat is their precise relationship?" *Desalination*, vol. 72, no. 3, pp. 275–292, 1989.

[62] A. Marandi, M. Polikarpus, and A. Jõeleht, "A new approach for describing the relationship between electrical conductivity and major anion concentration in natural waters," *Applied geochemistry*, vol. 38, pp. 103–109, 2013.

[63] D. Schulze-Makuch, "Longitudinal dispersivity data and implications for scaling behavior," *Groundwater*, vol. 43, no. 3, pp. 443–456, 2005.

[64] J. Wang, M. Han, A. B. Fuseni, D. Cao *et al.*, "Surfactant adsorption in surfactant-polymer flooding for carbonate reservoirs," in *SPE Middle East Oil & Gas Show and Conference*.   Society of Petroleum Engineers, 2015.

[65] S. Aramideh, R. Borgohain, P. K. Naik, C. T. Johnston, P. P. Vlachos, and A. M. Ardekani, "Multi-objective history matching of surfactant-polymer flooding," *Fuel*, vol. 228, pp. 418–428, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0016236118307014

[66] A. M. AlSofi, J. S. Liu, M. Han, and S. Aramco, "Numerical simulation of surfactant–polymer coreflooding experiments for carbonates," *Journal of Petroleum Science and Engineering*, vol. 111, pp. 184–196, 2013.

[67] A. Supee and A. K. Idris, "Effects of surfactant-polymer formulation and salinities variation towards oil recovery," *Arabian Journal for Science and Engineering*, vol. 39, no. 5, p. 42514260, Nov 2014.

[68] H. Bazargan, "An efficient polynomial chaos-based proxy model for history matching and uncertainty quantification of complex geological structures," Ph.D. dissertation, Heriot-Watt University, School of Petroleum Engineering, 2014.

[69] M. Williams, J. Keating, and M. Barghouty, "The stratigraphic method: A structured approach to history matching complex simulation models," *SPE Reservoir Evaluation and Engineering*, vol. 1, no. 02, pp. 169–176, Jan 1998.

[70] W. Kruger, "Determining areal permeability distribution by calculations," *Journal of Petroleum Technology*, vol. 13, no. 7, 1961.

[71] C. Romero, J. Carter, R. Zimmerman, and A. Gringarten, "Improved reservoir characterization through evolutionary computation," *Proceedings of SPE Annual Technical Conference and Exhibition*, 2000.

[72] G. Gavalas, P. Shah, and J. Seinfeld, "Reservoir history matching by bayesian estimation," *Society of Petroleum Engineers Journal*, vol. 16, no. 06, pp. 337–350, Jan 1976.

[73] P. Sharma, L. Durlofsky, and K. Aziz, "Efficient closed-loop production optimization under uncertainty," *Journal of Petroleum Technology*, vol. 57, no. 12, pp. 51–52, Jan 2005.

[74] Z. Tavassoli, J. N. Carter, and P. R. King, "Errors in history matching," *SPE Journal*, vol. 9, no. 03, pp. 352–361, Jan 2004.

[75] K. Tyler, T. Svanes, and S. Omdal, "Faster history matching and uncertainty in predicted production profiles with stochastic modeling," *Proceedings of SPE Annual Technical Conference and Exhibition*, 1993.

[76] W. Yeh, "Review of parameter identification in groundwater hydrology: the inverse problem," *Water Resources Research*, vol. Res. 22(2), pp. 95–108, February 1986.

[77] T. L. McLaughlin, D., "A reassessment of the groundwater inverse problem," *Water Resources Research*, vol. Res. 32(5), pp. 1131–1161, 1996.

[78] R. W. Rwechungura, M. Dadashpour, and J. Kleppe, "Advanced history matching techniques reviewed," *SPE Middle East Oil and Gas Show and Conference*, 2011.

[79] D. S. Oliver, A. C. Reynolds, and N. Liu, "Inverse theory for petroleum reservoir characterization and history matching," *Cambridge University Press*, 2008.

[80] L. Mohamed, M. A. Christie, and V. Demyanov, "Comparison of stochastic sampling algorithms for uncertainty quantification," *SPE Journal*, vol. 15, no. 01, pp. 31–38, Jan 2010.

[81] R. W. Schulze-Riegert and S. Selberg, "Event targeting model calibration used for history matching large simulation cases," *SPE Reservoir Simulation Symposium*, 2007.

[82] A. H. Elsheikh, M. D. Jackson, and T. C. Laforce, "Bayesian reservoir history matching considering model and parameter uncertainties," *Mathematical Geosciences*, vol. 44, no. 5, pp. 515–543, 2012.

[83] F. Zhang, J. A. Skjervheim, A. Reynolds, and D. Oliver, "Automatic history matching in a bayesian framework, example applications," *SPE Annual Technical Conference and Exhibition*, 2003.

[84] X. Ma, A. Datta-Gupta, and Y. Efendiev, "A multistage mcmc method with nonparametric error model for efficient uncertainty quantification in history matching," *SPE Annual Technical Conference and Exhibition*, 2008.

[85] L. Bonet-Cunha, D. Oliver, R. Redner, and A. Reynolds, "A hybrid markov chain monte carlo method for generating permeability fields conditioned to multiwell pressure data and prior information," *SPE Annual Technical Conference and Exhibition*, 1996.

[86] L. Mohamed, B. Calderhead, M. Filippone, M. Christie, and M. Girolami, "Population mcmc methods for history matching and uncertainty quantification," *Computational Geosciences*, vol. 16, no. 2, pp. 423–436, Oct 2011.

[87] A. Emerick and A. Reynolds, "Enkf-mcmc," *Proceedings of SPE EUROPEC/EAGE Annual Conference and Exhibition*, 2010.

[88] B. M. Negash, M. A. Ayoub, S. R. Jufar, and A. J. Robert, "History matching using proxy modeling and multiobjective optimizations," *Icipeg 2016*, pp. 3–16, 2017.

[89] J. He, J. Xie, X.-H. Wen, and W. Chen, "Improved proxy for history matching using proxy-for-data approach and reduced order modeling," *SPE Western Regional Meeting*, 2015.

[90] P. Sarma and J. Xie, "Efficient and robust uncertainty quantification in reservoir simulation with polynomial chaos expansions and non-intrusive spectral projection," *SPE Reservoir Simulation Symposium*, 2011.

[91] E. Fedutenko, C. Yang, C. Card, and L. X. Nghiem, "Time-dependent neural network based proxy modeling of sagd process," *SPE Heavy Oil Conference-Canada*, 2014.

[92] M. Sayyafzadeh, "History matching by online metamodeling," *SPE Reservoir Characterisation and Simulation Conference and Exhibition*, 2015.

[93] D. I. Zubarev, "Pros and cons of applying proxy-models as a substitute for full reservoir simulations," *SPE Annual Technical Conference and Exhibition*, 2009.

[94] L. Zerpa, N. Queipo, P. Salvador, and S. Jean-Louis, "An optimization methodology of alkaline-surfactant-polymer flooding processes using field scale numerical simulation and multiple surrogates," *Journal of Petroleum Science and Engineering*, pp. 197–208, 2005.

[95] S. Lucia, J. A. Paulson, R. Findeisen, and R. D. Braatz, "On stability of stochastic linear systems via polynomial chaos expansions," *2017 American Control Conference (ACC)*, 2017.

[96] F. S. Hover and M. S. Triantafyllou, "Application of polynomial chaos in stability and control," *Automatica*, vol. 42, no. 5, pp. 789–795, 2006.

[97] R. G. Ghanem and P. D. Spanos, *Stochastic finite elements: a spectral approach.* Springer, 1991.

[98] K. Sepahvand, S. Marburg, and H.-J. Hardtke, "Stochastic structural modal analysis involving uncertain parameters using generalized polynomial chaos expansion," *International Journal of Applied Mechanics*, vol. 03, no. 03, pp. 587–606, 2011.

[99] Z. Zhang, T. A. El-Moselhy, I. M. Elfadel, and L. Daniel, "Stochastic testing method for transistor-level uncertainty quantification based on generalized polynomial chaos," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 10, pp. 1533–1545, 2013.

[100] P. Sumant, H. Wu, A. Cangellaris, and N. Aluru, "Reduced-order models of finite element approximations of electromagnetic devices exhibiting statistical variability," *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 1, pp. 301–309, 2012.

[101] H. N. Najm, "Uncertainty quantification and polynomial chaos techniques in computational fluid dynamics," *Annual Review of Fluid Mechanics*, vol. 41, no. 1, pp. 35–52, 2009.

[102] N. Wiener, "The homogeneous chaos," *American Journal of Mathematics*, vol. 60, no. 4, p. 897, 1938.

[103] R. H. Cameron and W. T. Martin, "The orthogonal development of non-linear functionals in series of fourier-hermite functionals," *The Annals of Mathematics*, vol. 48, no. 2, p. 385, 1947.

[104] A. Siegel, T. Imamura, and W. C. Meecham, "Wienerhermite expansion in model turbulence in the late decay stage," *Journal of Mathematical Physics*, vol. 6, no. 5, pp. 707–721, 1965.

[105] W. C. Meecham and A. Siegel, "Wiener-hermite expansion in model turbulence at large Reynolds numbers," *Physics of Fluids*, vol. 7, no. 8, p. 1178, 1964.

[106] S. A. Orszag, "Dynamical properties of truncated wiener-hermite expansions," *Physics of Fluids*, vol. 10, no. 12, p. 2603, 1967.

[107] S. C. Crow and G. H. Canavan, "Relationship between a wienerhermite expansion and an energy cascade," *Journal of Fluid Mechanics*, vol. 41, no. 02, p. 387, 1970.

[108] H. Ogura, "Orthogonal functionals of the poisson process," *IEEE Transactions on Information Theory*, vol. 18, no. 4, pp. 473–481, 1972.

[109] D. Xiu and G. E. Karniadakis, "The wiener-askey polynomial chaos for stochastic differential equations," Jan 2003.

[110] X. Wan and G. E. Karniadakis, "Beyond wieneraskey expansions: Handling arbitrary pdfs," *Journal of Scientific Computing*, vol. 27, no. 1-3, pp. 455–464, 2005.

[111] A. Camacho, A. Talavera, A. A. Emerick, M. A. Pacheco, and J. Zanni, "Uncertainty quantification in reservoir simulation models with polynomial chaos expansions: Smolyak quadrature and regression method approach," *Journal of Petroleum Science and Engineering*, vol. 153, pp. 203–211, 2017.

[112] B. Sudret, "Global sensitivity analysis using polynomial chaos expansions," *Reliability Engineering and System Safety*, vol. 93, no. 7, pp. 964–979, 2008.

[113] O. Garcia-Cabrejo and A. Valocchi, "Global sensitivity analysis for multivariate output using polynomial chaos expansion," *Reliability Engineering and System Safety*, vol. 126, pp. 25–36, 2014.

[114] J. Rohmer, A. Loschetter, and D. Raucoules, "Global sensitivity analysis for supporting history matching of geomechanical reservoir models using satellite insar data," *Sensitivity Analysis in Earth Observation Modelling*, p. 145159, 2017.

[115] W. Ampomah, R. S. Balch, R. B. Grigg, B. Mcpherson, R. A. Will, S.-Y. Lee, Z. Dai, and F. Pan, "Co-optimization of co2-eor and storage processes in mature oil reservoirs," *Greenhouse Gases: Science and Technology*, vol. 7, no. 1, p. 128142, Feb 2016.

[116] R. Schulze-Riegert, J. Axmann, O. Haase, D. Rian, and Y.-L. You, "Optimization methods for history matching of complex reservoirs," *SPE Reservoir Simulation Symposium*, 2001.

[117] A. Wilson, "Proxy-based metamodeling optimization of gas-assisted-gravity-drainage process," *Journal of Petroleum Technology*, vol. 69, no. 10, pp. 92–94, Jan 2017.

[118] R. Patel, T. Jain, and J. Trivedi, "Polynomial-chaos-expansion based integrated dynamic modelling workflow for computationally efficient reservoir characterization: A field case study," *79th EAGE Conference and Exhibition 2017 - SPE EUROPEC*, Dec 2017.

[119] C. R. Xavier, E. P. D. Santos, V. D. F. Vieira, and R. W. D. Santos, "Genetic algorithm for the history matching problem," *Procedia Computer Science*, vol. 18, p. 946955, 2013.

[120] R. Schobi, B. Sudret, and J. Wiart, "Polynomial-chaos-based kriging," *International Journal for Uncertainty Quantification*, vol. 5, no. 2, pp. 171–193, 2015.

[121] I. Sobol, "Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates," *Mathematics and Computers in Simulation*, vol. 55, no. 1-3, pp. 271–280, 2001.

[122] M. Abily, O. Delestre, P. Gourbesville, N. Bertrand, C.-M. Duluc, and Y. Richet, "Global sensitivity analysis with 2d hydraulic codes: Application on uncertainties related to high-resolution topographic data," *Advances in Hydroinformatics Springer Water*, pp. 301–315, 2015.

[123] Q. Li, H. Xing, J. Liu, and X. Liu, "A review on hydraulic fracturing of unconventional reservoir," *Petroleum*, vol. 1, no. 1, p. 815, 2015.

[124] R. L. Johnson and C. W. Greenstreet, "Managing uncertainty related to hydraulic fracturing modeling in complex stress environments with pressure-dependent leakoff," *SPE Annual Technical Conference and Exhibition*, 2003.

[125] M. A. Bataweel, "Enhanced oil recovery in high salinity high temperature reservoir by chemical flooding," Ph.D. dissertation, University of Texas at Austin, Dec. 2011.

[126] D. L. Walker, "Experimental investigation of the effect of increasing the temperature on asp flooding," Ph.D. dissertation, University of Texas at Austin, Dec. 2011.

[127] P. T. Suniga, R. Fortenberry, and M. Delshad, "Observations of microemulsion viscosity for surfactant EOR processes," *SPE Improved Oil Recovery Conference*, 2016.

[128] G. A. Anderson, *Simulation of chemical flood enhanced oil recovery processes including the effects of reservoir wettability, Ph.D. thesis.* University of Texas at Austin, 2006.

[129] B. K. Paul and S. Moulik, "The viscosity behaviors of micro-emulsions: an overview," 2000.

[130] M. Delshad and G. Pope, "Comparison of the three-phase oil relative permeability models," *Transport in Porous Media*, vol. 4, no. 1, 1989.

[131] D. Schulze-Makuch, "Longitudinal dispersivity data and implications for scaling behavior," *Ground Water*, vol. 43, no. 3, p. 443456, 2005.

[132] A. M. Al-Sofi and M. J. Blunt, "The design and optimization of polymer flooding under uncertainty," *SPE Enhanced Oil Recovery Conference*, 2011.

[133] S. Marelli and B. Sudret, "UQLab user manual – polynomial chaos expansions," 2017, report UQLab-V1.0-104, doi:http://www.uqlab.com/download.

[134] G. Blatman and B. Sudret, "An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis," *Probabilistic Engineering Mechanics*, vol. 25, no. 2, pp. 183–197, 2010.

[135] S.-K. Choi, R. V. Grandhi, and R. A. Canfield, *Reliability-based structural design*. Springer, 2007.

[136] T. A. Mccourt, S. Hurter, B. Lawson, F. Zhou, B. Thompson, S. Tyson, and D. Donovan, "Uncertainty quantification of coal seam gas production prediction using polynomial chaos," *Journal of Petroleum Science and Engineering*, vol. 157, pp. 1148–1159, 2017.

[137] M. Hadigol and A. Doostan, "Least squares polynomial chaos expansion: A review of sampling strategies," *Computer Methods in Applied Mechanics and Engineering*, vol. 332, pp. 382–407, 2018.

[138] N. Fajraoui, S. Marelli, and B. Sudret, "Sequential design of experiment for sparse polynomial chaos expansions," *SIAM/ASA Journal on Uncertainty Quantification*, vol. 5, no. 1, pp. 1061–1085, 2017.

[139] C. Aistleitner, M. Hofer, and R. Tichy, "A central limit theorem for latin hypercube sampling with dependence and application to exotic basket option pricing," *International Journal of Theoretical and Applied Finance*, vol. 15, no. 07, 2012.

[140] M. Keramat and R. Kielbasa, "Latin hypercube sampling monte carlo estimation of average quality index for integrated circuits," *Analog Design Issues in Digital VLSI Circuits and Systems*, p. 131142, 1997.

[141] J.-H. Lee, Y.-D. Ko, I.-G. Yun, and K.-H. Han, "Comparison of latin hypercube sampling and simple random sampling applied to neural network modeling of hfo2thin film fabrication," *Transactions on Electrical and Electronic Materials*, vol. 7, no. 4, p. 210214, Jan 2006.

[142] V. Yaghoubi, S. Marelli, B. Sudret, and T. Abrahamsson, "Sparse polynomial chaos expansions of frequency response functions using stochastic frequency transformation," *Probabilistic Engineering Mechanics*, vol. 48, pp. 39–58, 2017.

[143] C. V. Mai and B. Sudret, "Surrogate models for oscillatory systems using sparse polynomial chaos expansions and stochastic time warping," *SIAM/ASA Journal on Uncertainty Quantification*, vol. 5, no. 1, pp. 540–571, 2017.

[144] G. Deman, K. Konakli, B. Sudret, J. Kerrou, P. Perrochet, and H. Benabder-rahmane, "Using sparse polynomial chaos expansions for the global sensitivity analysis of groundwater lifetime expectancy in a multi-layered hydrogeological model," *Reliability Engineering and System Safety*, vol. 147, pp. 156–169, 2016.

[145] K. Konakli and B. Sudret, "Polynomial meta-models with canonical low-rank approximations: Numerical insights and comparison to sparse polynomial chaos expansions," *Journal of Computational Physics*, vol. 321, pp. 1144–1169, 2016.

[146] C. V. Mai, M. D. Spiridonakos, E. N. Chatzi, and B. Sudret, "Surrogate modeling for stochastic dynamical systems by combining nonlinear autoregressive with exogenous input models and polynomial chaos expansions," *International Journal for Uncertainty Quantification*, vol. 6, no. 4, pp. 313–339, 2016.

[147] J. H. Holland, *Adaptation in Natural and Artificial Systems - An Introductory: Analysis with Applications to Biology, Control, and Artificial: Intelligence.* The University of Michigan Press, 1975.

[148] D. E. Goldberg, *Genetic algorithms in search, optimization, and machine learning.* Addison-Wesley, 2012.

[149] M. Sen, A. Datta-Gupta, P. Stoffa, L. Lake, and G. Pope, "Stochastic reservoir modeling using simulated annealing and genetic algorithm," *SPE Formation Evaluation*, vol. 10, no. 01, pp. 49–56, Jan 1995.

[150] H. Soleng, "Oil reservoir production forecasting with uncertainty estimation using genetic algorithms," *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*, 1999.

[151] C. Romero and J. Carter, "Using genetic algorithms for reservoir characterisation," *Journal of Petroleum Science and Engineering*, vol. 31, no. 2-4, pp. 113–123, 2001.

[152] M. Choudhary, S. Yoon, and B. Ludvigsen, "Application of global optimization methods for history matching andprobabilistic forecastingcase studies," *Proceedings of SPE Middle East Oil and Gas Show and Conference*, 2007.

[153] T. Reshma, K. V. Reddy, D. Pratap, M. Ahmedi, and V. Agilan, "Optimization of calibration parameters for an event based watershed model using genetic algorithm," *Water Resources Management*, vol. 29, no. 13, pp. 4589–4606, Jul 2015.

[154] L.-C. Chang, F.-J. Chang, K.-W. Wang, and S.-Y. Dai, "Constrained genetic algorithms for optimizing multi-use reservoir operation," *Journal of Hydrology*, vol. 390, no. 1-2, pp. 66–74, 2010.

[155] M. Shafii and F. D. Smedt, "Multi-objective calibration of a distributed hydrological model (wetspa) using a genetic algorithm," *Hydrology and Earth System Sciences*, vol. 13, no. 11, pp. 2137–2149, Dec 2009.

[156] J. A. Vrugt, H. V. Gupta, L. A. Bastidas, W. Bouten, and S. Sorooshian, "Effective and efficient algorithm for multiobjective optimization of hydrologic models," *Water Resources Research*, vol. 39, no. 8, 2003.

[157] B. Miller and D. E. Goldberg, *Genetic algorithms, tournament selection and the effects of noise*, 1995.

[158] A. Lipowski and D. Lipowska, "Roulette-wheel selection via stochastic acceptance," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 6, pp. 2193–2196, 2012.

[159] A. S. Uyar, "Experimental comparison of replacement strategies in steady state genetic algorithms for the dynamic mkp," *Lecture Notes in Computer Science Applications of Evolutinary Computing*, pp. 647–656, 2007.

[160] S. Picek, M. Golub, and D. Jakobovic, "Evaluation of crossover operator performance in genetic algorithms with binary representation," *Bio-Inspired Computing and Applications Lecture Notes in Computer Science*, pp. 223–230, 2012.

[161] S. M. Lim, A. B. M. Sultan, M. N. Sulaiman, A. Mustapha, and K. Y. Leong, "Crossover and mutation operators of genetic algorithms," *International Journal of Machine Learning and Computing*, vol. 7, no. 1, pp. 9–12, 2017.

[162] D. Savic, "Global and evolutionary optimization for water management problems," *Practical Hydroinformatics Water Science and Technology Library*, pp. 231–243, 2008.

[163] A. Efstratiadis and D. Koutsoyiannis, "Fitting hydrological models on multiple responses using the multiobjective evolutionary annealing-simplex approach," *Practical Hydroinformatics Water Science and Technology Library*, pp. 259–273, 2008.

[164] D. Krige, *A statistical approach to some basic mine valuation problems on the Witwatersrand.* Chemical, Metallurgical and Mining Society of South Africa, 1951.

[165] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning.* MIT Press, 2008.

[166] T. J. Santner, B. J. Williams, and W. I. Notz, *The design and analysis of computer experiments.* Springer, 2010.

[167] C. C. B. Cavalcante, C. Maschio, A. A. Santos, D. Schiozer, and A. Rocha, "History matching through dynamic decision-making," *Plos One*, vol. 12, no. 6, May 2017.

[168] T. Hastie, J. Taylor, R. Tibshirani, and G. Walther, "Forward stagewise regression and the monotone lasso," *Electronic Journal of Statistics*, vol. 1, p. 129, 2007.

[169] R. Tibshirani, "Regression shrinkage and selection via the lasso: a retrospective," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 73, no. 3, p. 27328, 2011.

[170] R. Tibshirani, I. Johnstone, T. Hastie, and B. Efron, "Least angle regression," *The Annals of Statistics*, vol. 32, no. 2, p. 407499, 2004.

[171] P. SWPU, "Fractured reservoir history matching improved based on artificial intelligent," vol. 2, p. 344360, 12 2016.

[172] H. Cui and M. G. Kelkar, "Automatic history matching of naturally fractured reservoirs and a case study," *SPE Western Regional Meeting*, 2005.

[173] Z. Zheng, "History matching and optimization using stochastic methods: Applications to chemical flooding," *PhD Thesis*, 2014.

[174] M. Alfi and S. A. Hosseini, "Integration of reservoir simulation, history matching, and 4D seismic for CO_2-EOR and storage at cranfield, mississippi, usa," *Fuel*, vol. 175, p. 116128, 2016.

[175] H. Class, L. Mahl, W. Ahmed, B. Norden, M. Khn, and T. Kempka, "Matching pressure measurements and observed CO_2 arrival times with static and dynamic modelling at the ketzin storage site," *Energy Procedia*, vol. 76, p. 623632, 2015.

[176] A. M. Alkhatib and P. R. King, "The use of the least squares probabilistic collocation method in decision making in the presence of uncertainty for chemical eor processes," *SPE Annual Technical Conference and Exhibition*, 2014.

[177] J. Zhang, M. Delshad, K. Sepehrnoori, and G. A. Pope, "An efficient reservoir-simulation approach to design and optimize improved oil-recovery-processes with distributed computing," *SPE Latin American and Caribbean Petroleum Engineering Conference*, 2005.

[178] M. J. Pyrcz and C. D. White, "Uncertainty in reservoir modeling," *Interpretation*, vol. 3, no. 2, 2015.

[179] F. Roggero and L. Hu, "Gradual deformation of continuous geostatistical models for history matching," *SPE Annual Technical Conference and Exhibition*, 1998.

[180] G. Evensen, J. Hove, H. Meisingset, E. Reiso, K. S. Seim, and . Espelid, "Using the enkf for assisted history matching of a north sea reservoir model," *SPE Reservoir Simulation Symposium*, 2007.

[181] Y. Gu and D. S. Oliver, "History matching of the punq-s3 reservoir model using the ensemble kalman filter," *SPE Annual Technical Conference and Exhibition*, 2004.

[182] G. Li and A. C. Reynolds, "Uncertainty quantification of reservoir performance predictions using a stochastic optimization algorithm," *Computational Geosciences*, vol. 15, no. 3, pp. 451–462, 2011.

[183] J. Caers, "Efficient gradual deformation using a streamline-based proxy method," *Journal of Petroleum Science and Engineering*, vol. 39, no. 1-2, p. 5783, 2003.

[184] L. Hu, "Gradual deformation and iterative calibration of gaussian-related stochastic models," *Math Geol*, pp. 87–108, 2000.

[185] J. Caers, "Comparing the gradual deformation with the probability perturbation method for solving inverse problems," *Mathematical Geology*, May 2007.

[186] S. Aanonsen, G. Naevdal, D. Oliver, A. Reynolds, and B. Valles, "The ensemble kalman filter in reservoir engineering - a review," *SPE J*, pp. 393–412, SPE J 2009 14(3).

[187] L. Heidari, V. Gervais, M. L. Ravalec, and H. Wackernagel, "History matching of petroleum reservoir models by the ensemble kalman filter and parameterization methods," *Computers Geosciences*, vol. 55, p. 8495, 2013.

[188] R. Tavakoli and A. C. Reynolds, "Monte carlo simulation of permeability fields and reservoir performance predictions with svd parameterization in rml compared with enkf," *Computational Geosciences*, vol. 15, no. 1, p. 99116, 2010.

[189] A. A. Emerick and A. C. Reynolds, "Combining the ensemble kalman filter with markov chain monte carlo for improved history matching and uncertainty characterization," *SPE Reservoir Simulation Symposium*, 2011.

[190] Á. Yustres, L. Asensio, J. Alonso, and V. Navarro, "A review of markov chain monte carlo and information theory tools for inverse problems in subsurface flow," *Computational Geosciences*, vol. 16, no. 1, pp. 1–20, 2012.

[191] L. Bonet-Cunha, D. Oliver, R. Redner, and A. Reynolds, "A hybrid markov chain monte carlo method for generating permeability fields conditioned to multiwell pressure data and prior information," *SPE Annual Technical Conference and Exhibition*, 1996.

[192] C. Maschio and D. J. Schiozer, "A new methodology for bayesian history matching using parallel interacting markov chain monte carlo," *Inverse Problems in Science and Engineering*, vol. 26, no. 4, p. 498529, 2017.

[193] L. Mohamed, B. Calderhead, M. Filippone, M. Christie, and M. Girolami, "Population mcmc methods for history matching and uncertainty quantification," *Computational Geosciences*, vol. 16, no. 2, p. 423436, Oct 2011.

[194] X. Ma, M. Al-Harbi, A. Datta-Gupta, and Y. Efendiev, "An efficient two-stage sampling method for uncertainty quantification in history matching geological models," *SPE Journal*, vol. 13, no. 01, p. 7787, Jan 2008.

[195] P. Naik, S. Aramideh, and A. M. Ardekani, "History matching of surfactant-polymer flooding using polynomial chaos expansion," *Journal of Petroleum Science and Engineering*, 2018.

[196] C. Maschio and D. J. Schiozer, "Bayesian history matching using artificial neural network and markov chain monte carlo," *Journal of Petroleum Science and Engineering*, vol. 123, p. 6271, 2014.

[197] S. Dachanuwattana, W. Yu, P. Zuloaga-Molero, and K. Sepehrnoori, "Application of assisted-history-matching workflow using proxy-based mcmc on a shale oil field case," *Journal of Petroleum Science and Engineering*, vol. 167, p. 316328, 2018.

[198] S. Dachanuwattana, J. Jin, P. Zuloaga-Molero, X. Li, Y. Xu, K. Sepehrnoori, W. Yu, and J. Miao, "Application of proxy-based mcmc and edfm to history match a vaca muerta shale oil well," *Fuel*, vol. 220, p. 490502, 2018.

[199] M. Wantawin, W. Yu, and K. Sepehrnoori, "An iterative work flow for history matching by use of design of experiment, response-surface methodology, and markov chain monte carlo algorithm applied to tight oil reservoirs," *SPE Reservoir Evaluation Engineering*, vol. 20, no. 03, p. 613626, Jan 2017.

[200] I. Bilionis and N. Zabaras, "Solution of inverse problems with limited forward solver evaluations: a bayesian perspective," *Inverse Problems*, vol. 30, no. 1, p. 015004, 2013.

[201] I. Bilionis, N. Zabaras, B. A. Konomi, and G. Lin, "Multi-output separable gaussian process: Towards an efficient, fully bayesian paradigm for uncertainty quantification," *Journal of Computational Physics*, vol. 241, pp. 212–239, 2013.

[202] I. Bilionis and N. Zabaras, "Multi-output local gaussian process regression: Applications to uncertainty quantification," *Journal of Computational Physics*, vol. 231, no. 17, pp. 5718–5746, 2012.

[203] H. Bazargan and M. Christie, "Bayesian model selection for complex geological structures using polynomial chaos proxy," *Computational Geosciences*, vol. 21, no. 3, p. 533551, Jan 2017.

[204] H. Bazargan, M. Christie, A. H. Elsheikh, and M. Ahmadi, "Surrogate accelerated sampling of reservoir models with complex structures using sparse polynomial chaos expansion," *Advances in Water Resources*, vol. 86, p. 385399, 2015.

[205] Y. Fu, J. Ding, H. Wang, and J. Wang, "Two-objective stochastic flow-shop scheduling with deteriorating and learning effect in industry 4.0-based manufacturing system," *Applied Soft Computing*, vol. 68, pp. 847–855, 2018.

[206] E. Zitzler, L. Thiele, M. Laumanns, C. M. Fonseca, and V. G. Da Fonseca, "Performance assessment of multiobjective optimizers: An analysis and review," *IEEE Transactions on evolutionary computation*, vol. 7, no. 2, pp. 117–132, 2003.

[207] D. R. Jones, M. Schonlau, and W. J. Welch, "Efficient global optimization of expensive black-box functions," *Journal of Global optimization*, vol. 13, no. 4, pp. 455–492, 1998.

[208] D. Huang, T. T. Allen, W. I. Notz, and N. Zeng, "Global optimization of stochastic black-box systems via sequential kriging meta-models," *Journal of global optimization*, vol. 34, no. 3, pp. 441–466, 2006.

[209] P. I. Frazier, W. B. Powell, and S. Dayanik, "A knowledge-gradient policy for sequential information collection," *SIAM Journal on Control and Optimization*, vol. 47, no. 5, pp. 2410–2439, 2008.

[210] P. Frazier, W. Powell, and S. Dayanik, "The knowledge-gradient policy for correlated normal beliefs," *Informs Journal on Computing*, vol. 21, no. 4, pp. 599–613, 2009.

[211] C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*, ser. Adaptive computation and machine learning.   Cambridge, MA: MIT Press, 2006.

[212] M. Emmerich, K. C. Giannakoglou, and . B. Naujoks, "Single-and multiobjective evolutionary optimization assisted by gaussian random field metamodels," *Evolutionary Computation, IEEE Transactions on*, vol. 10, no. 4, pp. 421–439, 2006.

[213] M. Emmerich, A. H. Deutz, and J. W. Klinkenberg, "Hypervolume-based expected improvement: Monotonicity properties and exact computation," in *Evolutionary Computation (CEC), 2011 IEEE Congress on.* IEEE, 2011, pp. 2147–2154.

[214] P. Pandita, I. Bilionis, J. Panchal, B. Gautham, A. Joshi, and P. Zagade, "Stochastic multiobjective optimization on a budget: Application to multipass wire drawing with quantified uncertainties," *International Journal for Uncertainty Quantification*, vol. 8, no. 3, 2018.

[215] G. J. Hirasaki, C. A. Miller, and M. Puerto, "Recent advances in surfactant eor," *SPE Annual Technical Conference and Exhibition*, 2008.

[216] G. Pope, B. Wang, and K. Tsaur, "A sensitivity study of micellar/polymer flooding," *Society of Petroleum Engineers Journal*, vol. 19, no. 06, p. 357368, Jan 1979.

[217] S. Solairaj, C. Britton, D. H. Kim, U. Weerasooriya, and G. A. Pope, "Measurement and analysis of surfactant retention," *SPE Improved Oil Recovery Symposium*, 2012.

[218] W. Weiss and R. Baldwin, "Planning and implementing a large-scale polymer flood," *Journal of Petroleum Technology*, vol. 37, no. 04, p. 720730, Jan 1985.

[219] D. Walker, C. Britton, D. H. Kim, S. Dufour, U. Weerasooriya, and G. A. Pope, "The impact of microemulsion viscosity on oil recovery," *SPE Improved Oil Recovery Symposium*, 2012.

[220] M. Tagavifar, S. Herath, U. P. Weerasooriya, K. Sepehrnoori, and G. Pope, "Measurement of microemulsion viscosity and its implications for chemical eor," *SPE Improved Oil Recovery Conference*, 2016.

[221] G. E. Box and G. C. Tiao, *Bayesian inference in statistical analysis.* John Wiley & Sons, 2011, vol. 40.

[222] D. S. Oliver, A. C. Reynolds, and N. Liu, *Inverse theory for petroleum reservoir characterization and history matching.* Cambridge University Press, 2008.

[223] M. L. Eaton and W. D. Sudderth, "Invariance of posterior distributions under reparametrization," *Sankhya A*, vol. 72, no. 1, pp. 101–118, 2010.

[224] W. K. Hastings, "Monte carlo sampling methods using markov chains and their applications," 1970.

[225] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions, and the bayesian restoration of images," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 721–741, 1984.

[226] H. Haario, E. Saksman, and J. Tamminen, "Adaptive proposal distribution for random walk metropolis algorithm," *Computational Statistics*, vol. 14, no. 3, pp. 375–396, 1999.

[227] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The journal of chemical physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[228] A. Patil, D. Huard, and C. J. Fonnesbeck, "Pymc: Bayesian stochastic modelling in python," *Journal of statistical software*, vol. 35, no. 4, p. 1, 2010.

[229] J. Geweke *et al.*, *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments.* Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991, vol. 196.

[230] F. Douarche, S. Da Veiga, M. Feraille, G. Enchéry, S. Touzani, and R. Barsalou, "Sensitivity analysis and optimization of surfactant-polymer flooding under uncertainties," *Oil & Gas Science and Technology–Revue dIFP Energies nouvelles*, vol. 69, no. 4, pp. 603–617, 2014.

[231] S. Le Van and B. H. Chon, "Chemical flooding in heavy-oil reservoirs: From technical investigation to optimization using response surface methodology," *Energies*, vol. 9, no. 9, p. 711, 2016.

[232] H. R. Jahangiri, D. Zhang *et al.*, "Optimization of the net present value of carbon dioxide sequestration and enhanced oil recovery," in *Offshore Technology Conference.* Offshore Technology Conference, 2011.

[233] H. Ahmed, A. A. Awotunde, A. S. Sultan, H. Y. Al-Yousef *et al.*, "Stochastic optimization approach to surfactant-polymer flooding," in *SPE/PAPG Pakistan Section Annual Technical Conference and Exhibition.* Society of Petroleum Engineers, 2017.

[234] P. Andonyadis, "Decision support for enhanced oil recovery projects," Ph.D. dissertation, 2010.

[235] P. Pandita, I. Bilionis, and J. Panchal, "Extending expected improvement for high-dimensional stochastic optimization of expensive black-box functions," *Journal of Mechanical Design*, vol. 138, no. 11, p. 111412, 2016.

[236] M. Stein, "Large sample properties of simulations using latin hypercube sampling," *Technometrics*, vol. 29, no. 2, pp. 143–151, 1987.

[237] G. Sun, G. Li, Z. Gong, X. Cui, X. Yang, and Q. Li, "Multiobjective robust optimization method for drawbead design in sheet metal forming," *Materials & Design*, vol. 31, no. 4, pp. 1917–1929, 2010.

[238] X. Gu, G. Sun, G. Li, L. Mao, and Q. Li, "A comparative study on multiobjective reliable and robust optimization for crashworthiness design of vehicle structure," *Structural and Multidisciplinary Optimization*, vol. 48, no. 3, pp. 669–684, 2013.

[239] A. Mondal, B. Mallick, Y. Efendiev, and A. Datta-Gupta, "Bayesian uncertainty quantification for subsurface inversion using a multiscale hierarchical model," *Technometrics*, vol. 56, no. 3, pp. 381–392, 2014.

[240] M. M. Khaninezhad and B. Jafarpour, "Sparse randomized maximum likelihood (sprml) for subsurface flow model calibration and uncertainty quantification," *Advances in water resources*, vol. 69, pp. 23–37, 2014.