Please fill in the name of the event you are preparing this manuscript for.	SPE Middle East Oil & Gas Show and Conference	
Please fill in your 6-digit SPE manuscript number.	SPE-204652-MS	
Please fill in your manuscript title.	Multi-fidelity Bayesian Approach for History Matching in Reservoir Simulation	
Please fill in your author name(s) and company affiliation.		
Given Name	Surname	Company
Ryan	Santoso	Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Kingdom of Saudi Arabia
Xupeng	He	Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Kingdom of Saudi Arabia
Marwa	Alsinan	EXPEC Advanced Research Center, Saudi Aramco, Kingdom of Saudi Arabia
Ruben	Figueroa Hernandez	Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Kingdom of Saudi Arabia
Hyung	Kwak	EXPEC Advanced Research Center, Saudi Aramco, Kingdom of Saudi Arabia
Hussein	Hoteit	Physical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Kingdom of Saudi Arabia

This template is provided to give authors a basic shell for preparing your manuscript for submittal to an SPE meeting or event. Styles have been included (Head1, Head2, Para, FigCaption, etc) to give you an idea of how your finalized paper will look before it is published by SPE. All manuscripts submitted to SPE will be extracted from this template and tagged into an XML format; SPE's standardized styles and fonts will be used when laying out the final manuscript. Links will be added to your manuscript for references, tables, and equations. Figures and tables should be placed directly after the first paragraph they are mentioned in. The technical content of your paper WILL NOT be changed. Please start your manuscript below.

Abstract

History matching is a critical step within the reservoir management process to synchronize the simulation model with the production data. The history-matched model can be used for planning optimum field development and performing optimization and uncertainty quantifications. We present a novel history matching workflow based on a Bayesian framework that accommodates subsurface uncertainties. Our workflow involves three different model resolutions within the Bayesian framework: 1) a coarse low-fidelity model to update the prior range, 2) a fine low-fidelity model to represent the high-fidelity model, and 3) a high-fidelity model to re-construct the real response. The low-fidelity model is constructed by a multivariate polynomial function, while the high-fidelity model is based on the reservoir simulation model. We firstly develop a coarse low-fidelity model using a two-level Design of Experiment (DoE), which aims to provide a better prior. We secondly use Latin Hypercube Sampling (LHS) to construct the fine low-fidelity model to be deployed in the Bayesian runs, where we use the Metropolis-Hastings algorithm. Finally, the posterior is fed into the high-fidelity model to evaluate the matching quality. This work demonstrates the importance of including uncertainties in history matching. Bayesian provides a robust framework to allow uncertainty quantification within the reservoir history matching. Under uniform prior, the convergence of the Bayesian is very sensitive to the parameter ranges. When the solution is far from the mean of the parameter ranges, the Bayesian introduces bias and deviates from the observed data. Our results show that updating the prior from the coarse low-fidelity model accelerates the Bayesian convergence and improves the matching convergence. Bayesian requires a huge number of runs to produce an accurate posterior.

Running the high-fidelity model multiple times is expensive. Our workflow tackles this problem by deploying a fine low-fidelity model to represent the high-fidelity model in the main runs. This fine low-fidelity model is fast to run, while it honors the physics and accuracy of the high-fidelity model. We also use ANOVA sensitivity analysis to measure the importance of each parameter. The ranking gives awareness to the significant ones that may contribute to the matching accuracy. We demonstrate our workflow for a geothermal reservoir with static and operational uncertainties. Our workflow produces accurate matching of thermal recovery factor and produced-enthalpy rate with physically-consistent posteriors. We present a novel workflow to account for uncertainty in reservoir history matching involving multi-resolution interaction. The proposed method is generic and can be readily applied within existing history-matching workflows in reservoir simulation.

Keywords: History matching, Adaptive Markov Chain Monte Carlo, Bayesian, Reduced-order

1. Introduction

Accurate capturing fluid-flow behavior in subsurface reservoirs is of importance for various environmental and energy applications, such as contaminant tracking in groundwater, carbon dioxide sequestration, and oil recovery. Aligning the simulation model with measured data (so-called *model calibration* or *history matching*) is required to achieve an accurate and physically-consistent forecast. The idea is to adjust uncertain parameters within the model (for instance, heterogeneity, fluid properties, porosity, and relative permeability) to reduce the mismatch between the model's prediction and measured data (AlAmeri, AlBreiki, and Geiger 2020; Elsheikh, Jackson, and Laforce 2012).

Model calibration has an inherently non-unique solution because of the high degree of freedom of adjusted parameters under limited measured data (Liu and Durlofsky 2020; Tang, Liu, and Durlofsky 2020). It also requires expensive computations to run the simulation model multiple times to evaluate the mismatch (Tang, Liu, and Durlofsky 2020; Elsheikh, Jackson, and Laforce 2012). Therefore, there are several challenges that should be addressed in performing model calibration: 1) cost of running the simulation model, 2) the means of including uncertainties during calibration, and 3) preserving the true physics.

The traditional approach relies on trial-error in adjusting the multipliers to minimize the mismatch between the model's forecast and measured data (Williams, Keating, and Barghouty 1998; Agarwal et al. 2000). It is inefficient, labor-intensive, and subject to personal interpretation (Awemo et al. 2014; Oliver, He, and Reynolds 1996).

Modern history matching methods use computerized algorithms to directly tune the parameters: Evolutionary Algorithm, Gradual Deformation, Neighborhood Algorithm, Randomized Maximum Likelihood, and Bayesian Markov Chain Monte Carlo (Kitanidis 1995; Schulze-Riegert et al. 2002; Caers 2003; Elsheikh, Jackson, and Laforce 2012). Bayesian Markov Chain Monte Carlo (MCMC) provides a rigorous framework to perform model calibration by considering uncertainties (Marzouk, Najm, and Rahn 2007; Elsheikh, Jackson, and Laforce 2012; Hamdi, Sousa, and Behmanesh 2017; Naik et al. 2019). It samples and develops the posterior distribution by utilizing the knowledge from the prior distribution through statistical samplings, such as Metropolis-Hastings, Reversible Jump, and Hamiltonian Monte Carlo (Zhang and Curtis 2020). It estimates the posterior distribution by calculating the likelihood, produced by running the forward model (Marzouk, Najm, and Rahn 2007). There are three critical issues influencing the precision, convergence, and efficiency: 1) proper selection of prior information, 2) optimum

calculation of the likelihood, and 3) optimum posterior samplings (Winkler 1967; Marzouk, Najm, and Rahn 2007; Mahendran et al. 2011; Jagalur-Mohan et al. 2018; Lemoine 2019).

Previous studies rarely deal with various prior distributions. Most of them assume informative priors (Gaussian distributions) (Marzouk, Najm, and Rahn 2007; Elsheikh, Jackson, and Laforce 2012; Hamdi, Sousa, and Behmanesh 2017). Naik et al. (2019) utilized symmetric uniform distribution as the priors. In geoscience, we frequently encounter sparse and limited measured data, therefore, informative prior is invalid. Furthermore, the inherently heterogeneous geological features also create a wide range of priors, in which the history matching solution might not be in the center of the range. Thus, we need to develop a new approach (within the Bayesian MCMC framework) to adapt to any prior condition, especially uninformative prior.

Every likelihood calculation requires a forward run. It is to quantify the mismatch between the model's forecast and measured data. Since Bayesian MCMC might have more than a hundred iterations and multiple chains to achieve convergence, likelihood calculation utilizing a high-fidelity model becomes expensive. We, therefore, require a low-fidelity model to represent the high-fidelity model. The commonly used low-fidelity models are Polynomial Chaos Expansion (PCE) and Gaussian Process Regression (GPR) (Marzouk, Najm, and Rahn 2007; Elsheikh, Jackson, and Laforce 2012; Naik et al. 2019). PCE provides an accurate approximation and guarantees physical consistency (Marzouk, Najm, and Rahn 2007; Blatman and Sudret 2011; Naik et al. 2019).

The commonly used statistical sampler for solving Bayesian MCMC are Metropolis-Hastings (MH), Adaptive Metropolis (AM), and Hamiltonian Monte Carlo (HMC) (Marzouk, Najm, and Rahn 2007; Mahendran et al. 2011; Zhang and Curtis 2020). The precision and convergence of the sampler are controlled by proper selection of hyperparameters, for instance, covariance scaling, tuning parameters, etc (Mahendran et al. 2011). Bayesian presents a rigorous optimization framework by treating the process as blackbox (Mahendran et al. 2011).

In this work, we develop a new Bayesian framework to efficiently perform the calibration with uncertainty quantification, optimally sample the posterior distributions, and flexibly deal with various prior distributions. We utilize multi-fidelity models to ensure the precision and physical consistency of the matching results. The idea is to firstly deploy coarse PCE, constructed using space-filling Design of Experiment (DoE). It is to update the initial prior ranges by exploring the response's space. We secondly utilize finer PCE, constructed using space-filling DoE with a bigger number of the training set. It is to represent the high-fidelity model and used for evaluating the likelihood. Thirdly, Bayesian-optimized Adaptive Metropolis is deployed to obtain the posterior distribution. Finally, the P50 (50% probability) values from the posteriors are fed into the high-fidelity model to check the physical consistency and accuracy. To our knowledge, this workflow is applied in subsurface flow problems for the first time.

We demonstrate the workflow in inter-well water injection in a vertically heterogeneous oil reservoir. The workflow provides predictions of time-dependent oil recovery factor (RF) and posterior distributions of vertical heterogeneity (Dykstra-Parson coefficient), oil viscosity, end-point oil relative permeability, and water injection rate. The correlation among parameters is identified through inferred covariance.

The outline of this paper is the following: In Section 2, we introduce the theories behind our proposed workflow. In Section 3, we show the problem of water injection in vertically heterogeneous oil reservoirs. In Section 4, we present the variability in subsurface water

injection based on California's projects. In Section 5, we explain the reservoir model, the physical phenomena, and initial priors. In Section 6, we present the workflow's results. Section 7 delivers sensitivity analysis on the number of the training set to construct PCE. In Section 8, we provide the discussions. Finally, we conclude the study in Section 9.

2. Proposed Workflow

2.1 Bayesian Markov Chain Monte Carlo (MCMC)

Let a forward problem be defined as follows

$$\mathbf{d} = \mathbf{G}(\mathbf{m}) + \boldsymbol{\varepsilon} \quad \boldsymbol{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}) \quad (1)$$

where $\mathbf{d}$ is a vector of observable quantities or data, $\mathbf{m}$ is model parameters, $\mathbf{G}$ is a forward model operator that yields predictions of $\mathbf{d}$ as a function of $\mathbf{m}$, $\boldsymbol{\varepsilon}$ describes the discrepancy between $\mathbf{d}$ and prediction of $\mathbf{G}(\mathbf{m})$. The discrepancy is assumed to be a normal distribution with zero mean value and covariance matrix, $\boldsymbol{\Sigma}$. The discrepancy might result from measurement error or model's inaccuracy (Wagner et al. 2019). $\mathbf{m}$ is an N_{in} -dimensional vector where $\mathbf{m} \in \Omega_{\mathbf{m}}$, $\mathbf{m} \in \{m_1, m_2, \dots, m_{N_{in}}\}$, and N_{in} is the input's dimension. $\mathbf{d}$ is an N_{out} -dimensional vector where $\mathbf{d} \in \{d_1, d_2, \dots, d_{N_{out}}\}$ and N_{out} is the output's dimension. $\mathbf{G}(\mathbf{m})$ produces N_{out} -dimensional vector of prediction. $\boldsymbol{\varepsilon}$ is an N_{out} -dimensional vector of error where $\boldsymbol{\varepsilon} \in \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_{N_{out}}\}$. $\boldsymbol{\Sigma}$ is a square matrix with $N_{out} \times N_{out}$ size.

The posterior distribution, $p(\mathbf{m} | \mathbf{d}, \mathbf{G})$, is evaluated using Bayes' theorem

$$p(\mathbf{m} | \mathbf{d}, \mathbf{G}) = \frac{p(\mathbf{d} | \mathbf{m}, \mathbf{G}) p(\mathbf{m} | \mathbf{G})}{\int_{\mathbf{m}} p(\mathbf{d} | \mathbf{m}, \mathbf{G}) p(\mathbf{m} | \mathbf{G}) d\mathbf{m}} \quad (2)$$

where $p(\mathbf{d} | \mathbf{m}, \mathbf{G})$ is the likelihood, $p(\mathbf{m} | \mathbf{G})$ is the prior distribution, and $\int_{\mathbf{m}} p(\mathbf{d} | \mathbf{m}, \mathbf{G}) p(\mathbf{m} | \mathbf{G}) d\mathbf{m}$ is the marginal likelihood. Since the discrepancy is generally unknown, the prior distribution becomes a joint distribution, $p^*(\mathbf{m} | \mathbf{G})$, between the model parameters and error's covariance

$$p^*(\mathbf{m} | \mathbf{G}) = p(\mathbf{m} | \mathbf{G}) \cdot p(\boldsymbol{\Sigma}) \quad (3)$$

where $p(\boldsymbol{\Sigma})$ is the prior of the error. The likelihood follows Gaussian formulation, expressed below

$$p(\mathbf{d} | \mathbf{m}, \mathbf{G}) = \prod_{i=1}^{N_{QoI}} \frac{1}{\sqrt{(2\pi)^{N_{out}} |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2}(\mathbf{d} - \mathbf{G}(\mathbf{m}))^T \boldsymbol{\Sigma}^{-1}(\mathbf{d} - \mathbf{G}(\mathbf{m}))\right) \quad (4)$$

where N_{QoI} is the number of Quantity of Interest (QoI).

The posterior distribution for each model parameter is calculated as follows

$$p(m_i) = \int_{\Omega_{\mathbf{m}_{\sim i}}} p(\mathbf{m} | \mathbf{d}, \mathbf{G}) d\mathbf{m}_{\sim i} \quad (5)$$

where $\mathbf{m}_{\sim i}$ refers to model parameters $\mathbf{m}$ excluding the i -th parameter m_i . The QoI prediction (posterior predictive) is written as follows

$$p(\mathbf{d}) = \int_{\Omega_{\mathbf{m}}} p(\mathbf{d} | \mathbf{m}, \mathbf{G}) p(\mathbf{m} | \mathbf{G}) d\mathbf{m} \quad (6)$$

The posterior distribution is constructed over the prior support $\Omega_{\mathbf{m}}$ through multiple iterations with transition probability, $p(\mathbf{m}^{(t+1)} | \mathbf{m}^{(t)})$. At iteration t , $\mathbf{m}^{(t)}$ is sampled from the priors and $\mathbf{m}^{(t+1)}$ is proposed for iteration $t+1$ following the below formulation

$$p(\mathbf{m}^{(t+1)}) = \int_{\Omega_{\mathbf{m}}} p(\mathbf{m}^{(t)}) p(\mathbf{m}^{(t+1)} | \mathbf{m}^{(t)}) d\mathbf{m}^{(t)} \quad (7)$$

The integration is solved using Monte Carlo approach.

2.2 Adaptive Metropolis algorithm

Adaptive Metropolis algorithm is proposed by Haario, Saksman, and Tamminen (2001). It is a modification of Metropolis-Hastings algorithm on the transition probability. The proposal process follows the random walk method in Metropolis-Hastings algorithm, in which the covariance is tuned using empirical parameter, s_d , and empirical covariance function (Gelman, Roberts, and Gilks 1994; Wagner et al. 2019). The mathematical expression is the following

$$\Sigma(t) = \begin{cases} \Sigma_0 & t \leq t_0 \\ s_d \cdot \Sigma(t-1) & t > t_0 \end{cases} \quad (8)$$

where t_0 refers to initial iteration and subscription 0 refers to the initial state. The empirical covariance function evaluates the proposed points from the previous realizations (up to iteration $t-1$) (Gelman, Roberts, and Gilks 1994; Wagner et al. 2019)

$$\Sigma(t) = \frac{1}{t} \left(\sum_{i=0}^t (\mathbf{m}^{(i)} - \bar{\mathbf{m}}^{(t)}) (\mathbf{m}^{(i)} - \bar{\mathbf{m}}^{(t)})^T \right), \text{ where } \bar{\mathbf{m}}^{(t)} = \frac{1}{t} \sum_{i=0}^t \mathbf{m}^{(i)} \quad (9)$$

The transition probability is then expressed as follows

$$p(\mathbf{m}^{(t+1)} | \mathbf{m}^{(t)}) = p(\mathbf{m}^{(t+1)} | \mathbf{m}^{(0)}, \dots, \mathbf{m}^{(t)}) = \square(\mathbf{m}^{(t+1)} | \mathbf{m}^{(t)}, \Sigma(t)) \quad (10)$$

The algorithm is extended to multiple chains to provide better initial seeds. Multiple chains can solve issues related to strongly correlated variables and multi-modal posterior (van Ravenzwaaij, Cassey, and Brown 2018; Martino 2018). Algorithm 1 shows the implementation of Multi-Chains Adaptive Metropolis algorithm. Each chain's realization can be independently evaluated.

Algorithm 1: Multi-Chain Adaptive Metropolis algorithm

Input: prior distribution of $\mathbf{m}$, number of iterations N_{iter} , number of chains N_{chains}

```

1  for  $w = 1, \dots, N_{chains}$  do
2      Initialize  $\mathbf{m}^{(0,w)}$ 
3      for  $k = 1, \dots, N_{iter}$  do
4          Calculate  $s_d \cdot \Sigma(k)$ 
5          Propose  $\mathbf{m}^{(*)} \sim p(\mathbf{m}^{(*,w)} | \mathbf{m}^{(0,w)}, \dots, \mathbf{m}^{(k-1,w)})$ 
6          Compute  $\alpha^{(k)} = \min \left\{ 1, \frac{p(\mathbf{m}^{(*,w)} | \mathbf{d}, \mathbf{G}) p(\mathbf{m}^{(0,w)}, \dots, \mathbf{m}^{(k-1,w)} | \mathbf{m}^{(*,w)})}{p(\mathbf{m}^{(k-1,w)} | \mathbf{d}, \mathbf{G}) p(\mathbf{m}^{(*,w)} | \mathbf{m}^{(0,w)}, \dots, \mathbf{m}^{(k-1,w)})} \right\}$ 
7          Generate  $u \sim U(0,1)$ 
8          If  $u \leq \alpha^{(k)}$  then  $\mathbf{m}^{(k,w)} \leftarrow \mathbf{m}^{(*,w)}$  else  $\mathbf{m}^{(k,w)} \leftarrow \mathbf{m}^{(k-1,w)}$ 
9      end
10 end

```

2.3 Bayesian optimization

The use of Bayesian optimization in MCMC aims to efficiently and accurately sample the posterior distribution. We would like to tackle the exhaustive trial-error hyperparameters tuning in the traditional approach.

The Bayesian optimization maximizes an objective function, defined as follows

$$\boldsymbol{\theta} = \max_{\boldsymbol{\theta} \in \square^{N_{opt}}} f(\boldsymbol{\theta}) \quad (11)$$

where $\boldsymbol{\theta}$ is the parameters to be optimized, N_{opt} is the number of optimized parameters, and $f(\boldsymbol{\theta})$ is objective function. Let the set of accumulated observations be written as follows

$$D_{1:N_{iter}} = \left\{ \boldsymbol{\theta}_{1:N_{iter}}, f(\boldsymbol{\theta}_{1:N_{iter}}) \right\} \quad (12)$$

where N_{iter} is the number of iterations within Bayesian optimization. The Bayesian formulation is the following

$$p(f | D_{1:N_{iter}}) = \frac{p(D_{1:N_{iter}} | f) p(f)}{\int_f p(D_{1:N_{iter}} | f) p(f) df} \quad (13)$$

where $p(f | D_{1:N_{iter}})$ is the posterior distribution of objective function after N_{iter} , $p(D_{1:N_{iter}} | f)$ is the likelihood after N_{iter} , $p(f)$ is the prior distribution, $\int_f p(D_{1:N_{iter}} | f) p(f) df$ is the marginal likelihood of objective function.

The posterior distribution is updated based on the *Expected Improvement* function

$$E[I] = \sigma(\theta) \cdot \left[\left(\frac{\mu(\theta) - f(\theta^+)}{\sigma(\theta)} \right) \Phi \left(\frac{\mu(\theta) - f(\theta^+)}{\sigma(\theta)} \right) + p \left(\frac{\mu(\theta) - f(\theta^+)}{\sigma(\theta)} \right) \right] \quad (14)$$

where $\sigma(\theta)$ is the standard deviation response of objective function as a function of θ , $\mu(\theta)$ is the mean response of objective function as a function of θ , $f(\theta^+)$ is the true maximum value of the objective function, and $\Phi(\cdot)$ is normal cumulative distribution. $\frac{\mu(\theta) - f(\theta^+)}{\sigma(\theta)}$ is analogous to z-value in normal distributions. The mean and standard deviation response of objective function is evaluated using Gaussian Process (GP).

GP of the posterior distribution is written as follows

$$f(\theta) \square GP(\mu(\theta), K(\theta, \theta')) \quad (15)$$

where $K(\theta, \theta')$ is the covariance function. The covariance matrix is a square matrix, written as follows

$$\mathbf{K}(\theta, \theta') = \begin{pmatrix} K(\theta_1, \theta_1) & \cdots & K(\theta_1, \theta_{N_{iter}}) \\ \vdots & \ddots & \vdots \\ K(\theta_{N_{iter}}, \theta_1) & \cdots & K(\theta_{N_{iter}}, \theta_{N_{iter}}) \end{pmatrix} \quad (16)$$

The size of the matrix grows as the number of iteration increases. The covariance function is squared exponential, written as follows (Brochu, Cora, and de Freitas 2010)

$$K(\theta, \theta') = \sigma^2(\theta) \cdot \exp \left(-\frac{1}{2w^2} \|\theta - \theta'\|^2 \right) \quad (17)$$

where w is the characteristic length controlling kernel width. At iteration t , the mean of the GP is calculated as follows

$$\mu(\theta^{(t)}) = \tilde{\mathbf{K}}^T \mathbf{K}^{-1} \mathbf{f}(\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(t-1)}) \quad (18)$$

$$\tilde{\mathbf{K}} = \begin{bmatrix} K(\theta^{(t)}, \theta^{(0)}) \\ \vdots \\ K(\theta^{(t)}, \theta^{(t-1)}) \end{bmatrix} \quad (19)$$

$$\mathbf{K} = \begin{bmatrix} K(\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(0)}) & \dots & K(\boldsymbol{\theta}^{(0)}, \boldsymbol{\theta}^{(t-1)}) \\ \vdots & \ddots & \vdots \\ K(\boldsymbol{\theta}^{(t-1)}, \boldsymbol{\theta}^{(0)}) & \dots & K(\boldsymbol{\theta}^{(t-1)}, \boldsymbol{\theta}^{(t-1)}) \end{bmatrix} \quad (20)$$

$$\mathbf{f} = \begin{bmatrix} f(\boldsymbol{\theta}^{(0)}) \\ \vdots \\ f(\boldsymbol{\theta}^{(t-1)}) \end{bmatrix} \quad (21)$$

The variance of the GP is calculated as follows

$$\sigma^2(\boldsymbol{\theta}^{(t)}) = K(\boldsymbol{\theta}^{(t)}, \boldsymbol{\theta}^{(t)}) - \tilde{\mathbf{K}}^T \mathbf{K}^{-1} \tilde{\mathbf{K}} \quad (22)$$

Algorithm 2 presents the implementation of Bayesian optimization. For Bayesian MCMC optimization, we propose a discrepancy between P50 prediction and measured data as the objective function.

Algorithm 2: Bayesian optimization algorithm

Input: initial seed $\boldsymbol{\theta}^{(0)}$, initial covariance $\tilde{\mathbf{K}}$, initial response $f(\boldsymbol{\theta}^{(0)})$, number of iterations

N_{iter}

```

1  for  $t = 1, \dots, N_{iter}$  do
2      Calculate the GP:  $\mu(\boldsymbol{\theta}^{(t)}), \sigma^2(\boldsymbol{\theta}^{(t)})$ 
3      Find  $\boldsymbol{\theta}^{(t)} = \arg \max_{\boldsymbol{\theta}} E[I]$ 
4      Calculate  $f(\boldsymbol{\theta}^{(t)})$ 
5      Augment the results  $D^{(t)} = \{D^{(t-1)}, (\boldsymbol{\theta}^{(t)}, f(\boldsymbol{\theta}^{(t)}))\}$ 
6  end
```

2.4 Polynomial Chaos Expansion (PCE)

Polynomial Chaos Expansion (PCE) serves as a low-fidelity model to represent the high-fidelity model. Its resolution can be controlled by the number of the training set, generated through space-filling DoE. The utilized space-filling DoE is Latin Hypercube.

PCE is a set of polynomial basis functions, written as follows

$$G(\boldsymbol{\xi}) = \sum_{\alpha \in A} g_{\alpha} \psi_{\alpha}(\boldsymbol{\xi}) \quad (23)$$

where $\psi_\alpha(\xi)$ is an orthogonal basis function, and ξ is a vector of random variables linked to model parameters. The subscript α denotes a multi-index scheme, where $\alpha \in A$ and

$$\alpha \equiv \alpha_q^l = \left\{ \begin{array}{c} \alpha_1^1, \alpha_1^2, \dots, \alpha_1^{(N_b+1)} \\ \alpha_2^1, \alpha_2^2, \dots, \alpha_2^{(N_b+1)} \\ \vdots \\ \alpha_{N_{opt}}^1, \alpha_{N_{opt}}^2, \dots, \alpha_{N_{opt}}^{(N_b+1)} \end{array} \right\}. \quad q \text{ is the index of model parameter and } l \text{ is the index of basis}$$

function. α returns the polynomial degree.

The basis function is truncated using total order scheme (Le Maitre and Knio 2010), expressed as

$$|\alpha| = \sum_{q=1}^{N_{opt}} \alpha_q^l \leq p \quad (24)$$

where p is the highest order within the polynomial basis. The number of basis function is expressed as follows

$$(N_b + 1) = \frac{(N_b + p)!}{N_b! p!} \quad (25)$$

The coefficient g is obtained using the Least Square method (Le Maitre and Knio 2010), which seeks to minimize the residual error, defined as

$$g = \arg \min_{g_\alpha \in \mathbb{R}} \frac{1}{N_s} \sum_{i=1}^{N_s} \left(\mathbf{d} - \sum_{\alpha \in A} g_\alpha \psi_\alpha(\xi) \right)^T \left(\mathbf{d} - \sum_{\alpha \in A} g_\alpha \psi_\alpha(\xi) \right) \quad (26)$$

where N_s is the number of the training set. The training set should be space-filling and free of outliers (Santoso, Hoteit, and Vahrenkamp 2019; Santoso, He, and Hoteit 2019; He, Santoso, and Hoteit 2020).

2.5 Proposed workflow

Figure 1 presents our proposed workflow for model calibration. The workflow has four important steps: 1) proper identification of model parameters, 2) updating initial priors, 3) evaluating posteriors with Bayesian-optimized Adaptive Metropolis and PCE, and 4) quality assessment.

Figure 1. Proposed workflow utilizing multi-fidelity models' interaction for model calibration. The black lines and black-line boxes denote the Bayesian MCMC loop. The red-line boxes denote the tool needed to construct the surrogate. The green-line boxes denote multi-resolution low-fidelity models used to ensure accurate and physically consistent history matching. The purple and brown line denote the loops for adjusting the models if the history matching result is still not accurate. P50 discrepancy reflects the error between P50 prediction and measured data.

Step 1 – model parameters identification: the model parameters are the uncertain parameters that control the variance and precision of the model's prediction. Each parameter is described with a distribution and variability (upper and lower bounds). The distribution and variability are obtained from data. This step is to set up initial priors.

Step 2 – updating initial priors: since the convergence and precision of Bayesian MCMC is strongly dependent on the prior selection. This step aims to update the initial prior to better bound the solution. We use Latin Hypercube to construct the coarse surrogate and assign new bounds based on the response's space exploration through the surrogate. We mainly focus on non-informative priors.

Figure 2 shows the detailed workflow for updating the priors. It is started by taking $N_{samples}$ using Latin Hypercube. We suggest 20 – 30 for the $N_{samples}$. The realizations from low-resolution PCE should bound the measured data. The new bounding lines should 1) cover at least 5 – 10% distance from the data, 2) not be too wide, and 3) put the data roughly in the center of the new bounds.

The bounding lines might not be unique. We should pick all bounding lines that follow the previous suggestions. As the number of samples increases, the number of bounding lines to be considered also increases. However, there will not be any overlapping bounds by increasing the number of samples.

Figure 2. Detailed process for updating initial priors.

The overall process to decide the optimum bounding lines is iterative. The optimum bounding lines are determined after the Bayesian MCMC run.

Step 3 – Bayesian-optimized Adaptive MCMC: the initial seeds for Bayesian MCMC is taken from the updated priors. For calculating the likelihood, the forward run utilizes high-resolution PCE to reduce the computational cost. The hyperparameters in the MCMC is tuned using Bayesian optimization. The objective function for the optimization is the discrepancy between P50 prediction and measured data. We utilize UQLab software for performing the Bayesian MCMC using Adaptive Metropolis (Marelli and Sudret 2014). The detailed implementation is in Algorithm 3.

Algorithm 3: Bayesian-optimized Adaptive MCMC

Input: prior distribution of $\mathbf{m}$, number of optimization steps N_{steps}

```

1  for  $t = 0, \dots, N_{steps}$  do
2      Run Adaptive Metropolis with  $\theta^{(t)} = \{N_{chains}, N_{iter}, s_d\}$ 
3      Calculate  $f(\theta^{(t)}) = \sqrt{\frac{1}{size(\mathbf{d})}(\mathbf{d} - \mathbf{P}_{50})^T(\mathbf{d} - \mathbf{P}_{50})}$ 
4      Augment the results  $D^{(t)} = \{D^{(t-1)}, (\theta^{(t)}, f(\theta^{(t)}))\}$ 
5      Calculate the GP:  $\mu(\theta^{(t)}), \sigma^2(\theta^{(t)})$ 
6      Find  $\theta^{(t)} = \arg \max_{\theta} E[I]$ 
7  end

```

Step 4 – quality assessment: the posterior distribution obtained from the MCMC is fed into the high-fidelity model to ensure accuracy and physical consistency. There are two inspections to be performed if the high-fidelity prediction is inconsistent: 1) revising the high-resolution PCE by increasing the number of training samples and 2) revising the matching parameters and re-checking the physics.

3. Problem formulation

We use 2-D vertical inter-well water injection in oil reservoirs as the case study. The water injection into oil reservoir (assuming constant temperature) is governed by two fundamental equations: mass conservation and momentum equation (Bird, Stewart, and Lightfoot 2006). The mass conservation in 2-D space is expressed below

$$\frac{\partial}{\partial t}(\phi \rho_{\eta} S_{\eta}) + \nabla \cdot (\rho_{\eta} u_{\eta}) = Q_{\eta} \quad \eta = \{o, w\} \quad (27)$$

where t is time, ϕ is the porosity of the reservoir, ρ is the phase density, S is the fluid phase saturation in the reservoir, u is the phase superficial velocity, and Q is the phase sink/source term. The subscription η denotes the phase where o is oil and w is water. Since we only consider two-phase oil and water, the phase saturation is constrained as follows

$$\sum_{i(\eta)}^2 S_i = 1 \quad (28)$$

The superficial velocity for each phase (momentum for each phase) is modeled using extended Darcy's law, written as follows

$$u_\eta = -\frac{\mathbf{k}k_{r\eta}}{\zeta_\eta} \nabla \Phi_\eta \quad (29)$$

where $\mathbf{k}$ is the absolute permeability of the reservoir, k_r is the phase relative permeability, ζ is the phase dynamic viscosity, and Φ is the phase potential, which is given by

$$\Phi_\eta = P_\eta - \rho_\eta \zeta z \quad (30)$$

where P is the phase pressure, ζ is the gravitational acceleration, and z is the phase hydrostatic height. The main objective is to calculate time-dependent oil recovery factor, written as follows

$$RF(t) = \frac{\int_0^t Q_{po} dt}{\phi \cdot RV \cdot B_o} \quad (31)$$

where Q_p is production rate, RV is reservoir bulk volume, and B is formation volume factor.

The density of the fluid is calculated using Equation of State (EoS) as a function of pressure and temperature (Firoozabadi 1999). The unknowns are pressure and saturation. In this work, the oil recovery factor is solved implicitly using CMG-IMEX simulator (CMG 2016).

3.1 Dimensionless variables

To be physically scalable for different static and dynamic conditions of a reservoir, we introduce a dimensionless time. The dimensionless time is defined as follows

$$t_D = \frac{Q_{iw} t}{\phi \cdot RV} \quad (32)$$

where Q_i is the injection rate. The dimensionless time represents the cumulative amount of injected water into the reservoir relative to the reservoir pore volume.

4. Variability in subsurface water injection

We compile data from waterflooding projects in California, USA (Van Wingen and Melkonian 1955). The dataset covers 55 projects in 21 oilfields in California. Five parameters are adopted to our modeling: permeability, porosity, thickness, oil viscosity, and water injection rate.

Figure 3. Data gathered from waterflooding projects in California, USA, reflecting the ranges of several parameters: permeability (a), porosity (b), thickness (c), viscosity (d), and injection rate (e). The data is adopted from Van Wingen and Melkonian (1955).

Figure 3 shows the statistics of the five adopted parameters. The permeability of water flooded fields in California spans from 40 mD to 2,500 mD, reflecting a wide range of variability due to

heterogeneity. The porosity varies between 0.15 and 0.4, while the thickness 20 and 380 *ft*. The oil viscosity is within the range of 1–600 *cP*, showing light to heavy oil. The average injection rate per rate is in the range of 14.4–10,300 *bbl / D*.

5. Simulation model

We utilize a model from the tenth SPE Comparative Solution Project as the high-fidelity physical model (Christie and Blunt 2001). The details about fluid properties, rock properties, and rock-fluid properties can be seen in Christie and Blunt (2001). We change the gridding from $100 \times 1 \times 20$ to $200 \times 1 \times 10$. The model resembles a layered-cake configuration where each layer is assumed to be homogeneous. This grid was verified and found to be stable and adequate for the simulation.

We account for the variation of permeability by assigning vertical heterogeneity into the ten layers (the assigned permeabilities are normalized) and taking P50 permeability (from data) as a multiplier (Santoso, Hoteit, and Vahrenkamp 2019; Santoso, Torrealba, and Hoteit 2020). The measure of vertical heterogeneity is through Dykstra-Parson coefficient (Jensen and Currie 1990; Santoso, Hoteit, and Vahrenkamp 2019; Santoso, Torrealba, and Hoteit 2020), expressed as follows

$$V_{DP} = \frac{k_{0.5} - k_{0.16}}{k_{0.5}} \quad (33)$$

where subscription 0.5 denotes 50th percentile, and 0.16 denotes 16th percentile from the Cumulative Distribution Function (CDF). The variation is in the range of 0.5–0.9, reflecting homogeneous to the heterogeneous case.

We vary the end-point oil relative permeability to reflect the degree of oil wetting in the reservoir. The variation is in the range of 0.6–0.9. The upper bound comes from the original value in Christie and Blunt (2001), while the lower bound is to show stronger oil-wet.

In summary, four parameters are varied in the model: vertical heterogeneity, oil viscosity, end-point oil relative permeability, and water injection rate. The permeability multiplier is fixed at 272.5 *mD* (P50 permeability). The porosity is constant at 0.25 (P50 porosity). The thickness is also constant at 137.5 *ft* (P50 thickness), therefore, each layer has 13.75 *ft* thickness. The water is injected at a constant rate and the bottom-hole pressure is kept constant at 3,000 *psi*. The model in varying vertical heterogeneity is in Figure 4. We show the base-case simulation under the following setup: $V_{DP} = \{0.5, 0.7, 0.9\}$, $\zeta_o = 33.8$ *cP*, $k_{ro} = 0.9$, and $Q_{iw} = 786.5$ *bbl/D*.

Figure 4. Variation in vertical heterogeneity: $V_{DP} = 0.5$ (first column), $V_{DP} = 0.7$ (second column), and $V_{DP} = 0.9$ (third column).

Figure 5 shows the injected-water front propagation under different vertical heterogeneity. The typical flow occurring in the reservoir is viscous-dominated. The flow is mainly controlled by the mobility ratio and injection rate (Lenormand, Touboul, and Zarcone 1988). The water tends to channel through high-permeability layers and bypass the oil. Since the water viscosity is below the oil viscosity (In our dataset, the lowest oil viscosity is 1 cP), unstable displacement occurs (Lenormand, Touboul, and Zarcone 1988). The unstable displacement develops a fingering profile, which causes early water breakthrough at the producer. Early water breakthrough reduces oil production.

There is no significant difference between oil density and water density. Therefore, gravitational force does not significantly act on the fluids and does not cause segregation (Huppert and Neufeld 2014).

Figure 5. Water saturation profile at different times showing the injected-water front propagation. The first column is for $V_{DP} = 0.5$, second for $V_{DP} = 0.7$, and third for $V_{DP} = 0.9$. The first row is for $t_D = 0.1$, second for $t_D = 1.0$, and third for $t_D = 3.0$.

Figure 6a shows the bottom-hole pressure as a function of dimensionless time at both producer and injector under different vertical heterogeneity. The bottom-hole pressure at the producer is constant at 3,000 psi over time, reflecting the constraint. The bottom-hole pressure at the injector jumps from 5,000 psi (initial pressure), then, gradually declines. The spike in the pressure shows

the process of overcoming inlet resistances, such as entry pressure. Once the displacement pathways are established, the pressure starts to go down gradually (going to equilibrium condition). The water breakthrough time is observed through the change in the trend of the declining-pressure profile. The bottom-hole injection pressure at homogeneous conditions is higher than at heterogeneous conditions.

Figure 6. Simulation results for several scenarios showing bottom-hole pressure vs dimensionless time at both injector (continuous line) and producer (dashed-line) (a), oil rate vs dimensionless time (b), oil recovery factor vs dimensionless time (c), water-cut vs dimensionless time (d).

Figure 6b shows the oil rate profile over time under different heterogeneity. At homogeneous conditions, the injected-water is distributed uniformly to all layers, therefore, it is efficiently swept the oil. Efficient sweeping leads to high oil rate and oil recovery factor (see Figure 6c).

Figure 6d presents the water-cut profile. At heterogeneous conditions, there is an early water breakthrough. It is denoted by an early increase in the water-cut profile. Early water breakthrough reduces the oil production, therefore, the water-cut is the highest, reflecting enormous water production.

6. Results

We adopt field data from Huntington Beach Jones, USA for the reference case (Van Wingen and Melkonian 1955). The setup is the following: $V_{DP} = 0.70$, $\zeta_o = 130 \text{ cP}$, $k_{ro} = 0.65$, and $Q_{iw} = 1,008 \text{ bbl/D}$.

Figure 7 show the updated bounding lines obtained from 25 realizations of Latin Hypercube. We explore 20, 25, and 30 realizations of Latin Hypercube to find the proper bounding lines (see Figure A-1). We accommodate all possible bounding lines and find the optimum one through the iterative process (iterative process from Step 2 to Step 4 in Section 2). The new bounds are the following: $V_{DP} = 0.63 - 0.89$, $\zeta_o = 110.12 - 199.16 \text{ cP}$, $k_{ro} = 0.64 - 0.77$, and $Q_{iw} = 362.15 - 2,272 \text{ bbl/D}$.

Figure 7. Updated bounding lines (b) from 25 realizations of Latin Hypercube. The realizations are random.

We perform sensitivity analysis on the number of the training set to find the optimum high-resolution PCE representation. The sensitivity analysis is presented in Section 7. Optimum PCE is achieved when 1) it has high accuracy with low rank/low degree and 2) it has high predictability under the same testing set. The optimum high-resolution PCE has 1,000 training set, shown in Figure 8.

Figure 8. The training and testing quality of the High-resolution PCE: RF training accuracy (a), RF prediction accuracy (b), and average RMSE (c). The (a) and (b) plots only show some samples.

We perform Bayesian-optimized Adaptive MCMC with prior covariance scaling 0.0001–1.0, number of parallel chains 1–500, and length of chains 1–100,000. We burn-in 50% of the results. The sensitivities are in Section 7. The optimum scenario is the following: $N_{chains} = 395$, $N_{iter} = 91,412$, and $s_d = 0.26$. The optimum scenario produces: 1) small discrepancy between P50 prediction and measured data and 2) narrow variance. The RF posterior prediction and the variance from the Bayesian runs are in Figure 9. The posterior distributions and statistics of matching parameters are in Figure 10. The comparison of reference permeability and matched permeability is in Figure 11.

Figure 9. Posterior prediction of RF (a) and variance (b) from Bayesian MCMC runs. The variance distribution is smoothed using Kernel Density Estimator (KDE).

Figure 10. Prior, updated prior, and posterior distribution for heterogeneity (a), oil viscosity (b), end-point oil relative permeability (c), and injection rate (d). Posterior distributions are smoothed using Kernel Density Estimator (KDE).

Figure 11. Correlation among matching parameters. The figures in the main diagonal show the posterior distribution of each parameter. The ones in the lower triangle show the covariance: the red line is associated with 10% probability, the blue line 50%, and the green line 90%. The ones in the upper triangle are the correlation coefficient. The X and Y scale in the covariance plot span between -1 and 1.

(a)

(b)

Figure 12. Comparison of high-fidelity prediction using Bayesian's posteriors and the data (a) and permeability realization from P50 posterior (b).

We take the P50 value from each posterior distribution and re-calculate the high-fidelity response. Figure 12 shows the comparison using the high-fidelity model and permeability realization from P50 posterior results. The P50 prediction from Bayesian run agrees with the data. This result shows that the Bayesian results are accurate.

7. Sensitivity analysis

We perform sensitivity analysis on the number of the training set (generated through Latin Hypercube) for constructing the high-resolution PCE. It is critical since the model represents the high-fidelity model in the Bayesian runs.

We consider 5 – 1,500, the number of training set for the construction. Figure 13 shows the training accuracy and predictability of each selection.

Figure 13. The effect of changing the number of training set towards training RMSE (a) and prediction RMSE (b). The prediction is calculated using 1,500 blind sets.

Chaotic trend is observed in the training RMSE by increasing the number of the training set. It occurs due to changing of the polynomial degree to fit the training data. However, a decreasing trend is observed in prediction RMSE by increasing the number of the training set. Since there is no significant difference in prediction RMSE between 1,000 and 1,500 training sets, we pick 1,000 training sets as the optimum one.

8. Discussions

Increasing the number of realizations for updating the priors means increasing the chance to find the proper bounding lines. Since Latin Hypercube always guarantees space-filling samples, the bounding lines become narrower (respectively to the data) by increasing the number of realizations.

There is non-uniqueness in selecting the proper bounding lines (see Figure A-1). Since the realizations of Latin Hypercube is random, there are many possible bounding lines. At this step, we need to utilize all the bounding lines and decide the optimum one after going through Step 4

(in *Proposed Workflow* section). The uniqueness of the objective functions determines the number of possible bounding lines to be executed.

The accuracy of high-resolution PCE determines the convergence and reliability of the Bayesian MCMC results. The accuracy affects the consistency between P50 Bayesian results on the high-fidelity model and the data. Furthermore, we want to emphasize that predictability of the PCE is more significant than training accuracy. The predictability shows the reliability of the PCE in predicting blind sets. There is a saturation level of predictability by increasing number of training set. As it can be seen in Figure 13b, 1,000 and 1,500 training set does not give significant difference in predictability.

The optimized Bayesian MCMC provides an efficient and reliable matching result. Bayesian optimization is a robust approach to optimize the posterior construction in Bayesian MCMC. All hyperparameters within the Bayesian MCMC (covariance scaling, number of chains, and chain's length) should be tuned simultaneously (Mahendran et al. 2011). The optimized covariance scaling ensures fast convergence (Mahendran et al. 2011). The optimized number of chains and chain's length reduce the computational cost (Mahendran et al. 2011; Elsheikh, Jackson, and Laforce 2012).

The correlation/covariance in Figure 11 captures the physical processes in the model. It shows the correlation among parameters governed by the RF results. The P50 Bayesian posterior on the high-fidelity model shows accurate matching (see Figure 12a). An accurate P50 prediction on a high-fidelity model ensures physical consistency.

9. Conclusions

We present a new workflow to efficiently perform history matching considering uncertainty. The workflow uses a Bayesian framework for addressing uncertainties during the matching process and PCE for reducing the computational cost. The Bayesian MCMC is optimized through Bayesian optimization to ensure convergence and reliability.

Updating the initial priors improves the precision of the Bayesian results. The selection of bounding lines is non-unique, therefore, the optimum one is later found after going through the quality assessment step in the workflow.

The accuracy of the high-resolution PCE determines the quality of the matching. Predictability of the PCE is more important than training accuracy. We found saturation of PCE's predictability by increasing the number of the training set for the construction.

The optimized Bayesian MCMC provides efficient and robust matching. The optimized covariance scaling ensures fast convergence, while the optimized number of chains and chain's length reduce the computational cost.

The P50 Bayesian results prediction on the high-fidelity model ensures physical consistency and robustness of the matching. The covariance/correlation among parameters reflects the physical processes within the model.

Acknowledgement

We would like to thank CMG Ltd. for providing the IMEX academic license and UQLab for the software license.

References

- Agarwal, B., H. Hermansen, J. E. Sylte, and L. K. Thomas. 2000. "Reservoir Characterization of Ekofisk Field: A Giant, Fractured Chalk Reservoir in the Norwegian North Sea-History Match." *SPE Reservoir Evaluation and Engineering* 3 (6): 534–43. <https://doi.org/10.2118/68096-PA>.
- AlAmeri, Saeeda, Mohamed AlBreiki, and Sebastian Geiger. 2020. "History Matching under Geological Constraints Coupled with Multiobjective Optimization to Optimize MWAG Performance: A Case Study in a Giant Onshore Carbonate Reservoir in the Middle East." *SPE Reservoir Evaluation and Engineering* 23 (2): 534–50. <https://doi.org/10.2118/196715-PA>.
- Awemo, K N, I. Ajala, C. A. Schwietzer, L. Ganzer, R. Schulze-Riegert, and H. Almualim. 2014. "P05 Assisted History Matching with Application of Adjoint Method Sensitivity Computation: Case Study North German Basin Oilfield." In *2nd EAGE Integrated Reservoir Modelling Conference*. Dubai, UAE.
- Bird, R. B., W. E. Stewart, and E. N. Lightfoot. 2006. *Transport Phenomena*. 2nd ed. New York: Wiley. <https://www.wiley.com/en-us/Transport+Phenomena%2C+Revised+2nd+Edition-p-9780470115398>.
- Blatman, Géraud, and Bruno Sudret. 2011. "Adaptive Sparse Polynomial Chaos Expansion Based on Least Angle Regression." *Journal of Computational Physics* 230 (6): 2345–67. <https://doi.org/10.1016/j.jcp.2010.12.021>.
- Brochu, Eric, Vlad M. Cora, and Nando de Freitas. 2010. "A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning," December. <http://arxiv.org/abs/1012.2599>.
- Caers, Jef. 2003. "Efficient Gradual Deformation Using a Streamline-Based Proxy Method." *Journal of Petroleum Science and Engineering* 39 (1–2): 57–83. [https://doi.org/10.1016/S0920-4105\(03\)00040-8](https://doi.org/10.1016/S0920-4105(03)00040-8).
- Christie, M. A., and M. J. Blunt. 2001. "Tenth SPE Comparative Solution Project: A Comparison of Upscaling Techniques." *SPE Reservoir Evaluation and Engineering* 4 (4): 308–16. <https://doi.org/10.2118/72469-pa>.
- CMG. 2016. "IMEX User Guide." Calgary, Canada.
- Elsheikh, A. H., M. D. Jackson, and T. C. Laforce. 2012. "Bayesian Reservoir History Matching Considering Model and Parameter Uncertainties." *Mathematical Geosciences* 44 (5): 515–43. <https://doi.org/10.1007/s11004-012-9397-2>.
- Firoozabadi, Abbas. 1999. *Thermodynamics of Hydrocarbon Reservoirs*. McGraw-Hill. <https://agris.fao.org/agris-search/search.do?recordID=US201300029874>.
- Gelman, A., G. O. Roberts, and W. R. Gilks. 1994. "Efficient Metropolis Jumping Rules." In *Bayesian Statistics 5, Proceedings of the 5th Valencia International Meeting*, 599–607. Oxford University Press.
- Haario, Heikki, Eero Saksman, and Johanna Tamminen. 2001. "An Adaptive Metropolis Algorithm." *Bernoulli* 7 (2): 223–42. <https://doi.org/10.2307/3318737>.
- Hamdi, Hamidreza, Mario Costa Sousa, and Hamid Behmanesh. 2017. "Bayesian History-Matching and Probabilistic Forecasting for Tight and Shale Wells." In *SPE Unconventional*

- Resources Conference*, 2:969–81. Society of Petroleum Engineers. <https://doi.org/10.2118/185082-ms>.
- He, Xupeng, Ryan Santoso, and Hussein Hoteit. 2020. “Application of Machine-Learning to Construct Equivalent Continuum Models from High-Resolution Discrete-Fracture Models.” In *International Petroleum Technology Conference*. Society of Petroleum Engineers (SPE). <https://doi.org/10.2523/iptc-20040-ms>.
- Huppert, Herbert E., and Jerome A. Neufeld. 2014. “The Fluid Mechanics of Carbon Dioxide Sequestration.” *Annual Review of Fluid Mechanics* 46 (1): 255–72. <https://doi.org/10.1146/annurev-fluid-011212-140627>.
- Jagalur- Mohan, Jayanth, Birendra Jha, Zheng Wang, Ruben Juanes, and Youssef Marzouk. 2018. “Inferring Fault Frictional and Reservoir Hydraulic Properties From Injection- Induced Seismicity.” *Geophysical Research Letters* 45 (3): 1313–20. <https://doi.org/10.1002/2017GL075925>.
- Jensen, Jerry L., and Iain D. Currie. 1990. “New Method for Estimating the Dykstra-Parsons Coefficient to Characterize Reservoir Heterogeneity.” *SPE Reservoir Engineering* 5 (3): 369–74. <https://doi.org/10.2118/17364-pa>.
- Kitanidis, Peter K. 1995. “Quasi- Linear Geostatistical Theory for Inversing.” *Water Resources Research* 31 (10): 2411–19. <https://doi.org/10.1029/95WR01945>.
- Lemoine, Nathan P. 2019. “Moving beyond Noninformative Priors: Why and How to Choose Weakly Informative Priors in Bayesian Analyses.” *Oikos* 128 (7): 912–28. <https://doi.org/10.1111/oik.05985>.
- Lenormand, Roland, Eric Touboul, and Cesar Zarcone. 1988. “Numerical Models and Experiments on Immiscible Displacements in Porous Media.” *Journal of Fluid Mechanics* 189: 165–87. <https://doi.org/10.1017/S0022112088000953>.
- Liu, Yimin, and Louis J. Durlofsky. 2020. “Multilevel Strategies and Geological Parameterizations for History Matching Complex Reservoir Models.” *SPE Journal* 25 (1): 81–104. <https://doi.org/10.2118/193895-PA>.
- Mahendran, Nimalan, Ziyu Wang, Firas Hamze, and Nando de Freitas. 2011. “Bayesian Optimization for Adaptive MCMC,” October. <http://arxiv.org/abs/1110.6497>.
- Maitre, O. P. Le, and O. M. Knio. 2010. *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*. Media. New York: Springer. https://doi.org/10.1007/978-90-481-3520-2_1.
- Marelli, Stefano, and Bruno Sudret. 2014. “UQLab: A Framework for Uncertainty Quantification in MATLAB.” In *The 2nd International Conference on Vulnerability and Risk Analysis and Management (ICVRAM 2014)*, 2554–63. <https://doi.org/10.1061/9780784413609.257>.
- Martino, Luca. 2018. “A Review of Multiple Try MCMC Algorithms for Signal Processing.” *Digital Signal Processing* 75 (April): 134–52. <https://doi.org/10.1016/J.DSP.2018.01.004>.
- Marzouk, Youssef M., Habib N. Najm, and Larry A. Rahn. 2007. “Stochastic Spectral Methods for Efficient Bayesian Solution of Inverse Problems.” *Journal of Computational Physics* 224 (2): 560–86. <https://doi.org/10.1016/j.jcp.2006.10.010>.
- Naik, Pratik, Piyush Pandita, Soroush Aramideh, Ilias Bilonis, and Arezoo M. Ardekani. 2019. “Bayesian Model Calibration and Optimization of Surfactant-Polymer Flooding.” *Computational Geosciences* 23 (5): 981–96. <https://doi.org/10.1007/s10596-019-09858-z>.
- Oliver, D. S., N. He, and A. C. Reynolds. 1996. “Conditioning Permeability Fields to Pressure Data.” In *5th European Conference on Mathematics of Oil Recovery*, 259–70. Leoben, Austria: European Association of Geoscientist and Engineers.
- Ravenzwaaij, Don van, Pete Cassey, and Scott D. Brown. 2018. “A Simple Introduction to Markov Chain Monte-Carlo Sampling.” *Psychonomic Bulletin and Review* 25 (1): 143–54. <https://doi.org/10.3758/s13423-016-1015-8>.
- Santoso, Ryan, Xupeng He, and Hussein Hoteit. 2019. “Application of Machine-Learning to

- Construct Simulation Models from High-Resolution Fractured Formation.” In *Abu Dhabi International Petroleum Exhibition and Conference*. Society of Petroleum Engineers (SPE). <https://doi.org/10.2118/197439-ms>.
- Santoso, Ryan, Hussein Hoteit, and Volker Vahrenkamp. 2019. “Optimization of Energy Recovery from Geothermal Reservoirs Undergoing Re-Injection: Conceptual Application in Saudi Arabia.” In *SPE Middle East Oil and Gas Show and Conference*. Society of Petroleum Engineers (SPE). <https://doi.org/10.2118/195155-ms>.
- Santoso, Ryan, Victor Torrealba, and Hussein Hoteit. 2020. “Investigation of an Improved Polymer Flooding Scheme by Compositionally-Tuned Slugs.” *Processes* 8 (2): 197. <https://doi.org/10.3390/pr8020197>.
- Schulze-Riegert, R. W., J. K. Axmann, O. Haase, D. T. Rian, and Y. L. You. 2002. “Evolutionary Algorithms Applied to History Matching of Complex Reservoirs.” *SPE Reservoir Evaluation and Engineering* 5 (2): 163–73. <https://doi.org/10.2118/77301-PA>.
- Tang, Meng, Yimin Liu, and Louis J. Durlofsky. 2020. “A Deep-Learning-Based Surrogate Model for Data Assimilation in Dynamic Subsurface Flow Problems.” *Journal of Computational Physics* 413 (July): 109456. <https://doi.org/10.1016/j.jcp.2020.109456>.
- Wagner, P. R., J. Nagel, S. Marelli, and B. Sudret. 2019. “UQLab User Manual -- Bayesian Inversion for Model Calibration and Validation.”
- Williams, M. A., J. F. Keating, and M. F. Barghouty. 1998. “The Stratigraphic Method: A Structured Approach to History-Matching Complex Simulation Models.” *SPE Reservoir Engineering (Society of Petroleum Engineers)* 1 (2): 169–76. <https://doi.org/10.2118/38014-pa>.
- Wingen, N. Van, and B. L. Melkonian. 1955. “Review of California Waterflooding Operations.” *Journal of Petroleum Technology* 13 (02): 130–33. <https://doi.org/10.2118/1599-g-pa>.
- Winkler, Robert L. 1967. “The Assessment of Prior Distributions in Bayesian Analysis.” *Journal of the American Statistical Association* 62 (319): 776. <https://doi.org/10.2307/2283671>.
- Zhang, Xin, and Andrew Curtis. 2020. “Seismic Tomography Using Variational Inference Methods.” *Journal of Geophysical Research: Solid Earth* 125 (4). <https://doi.org/10.1029/2019jb018589>.

Appendix A

We show 20, 25, and 30 realizations of Latin Hypercube in Section 6 in Figure A-1.

Figure A-1. Latin Hypercube realizations for updating the priors: 20 (a), 25 (b), and 30 (c). The lines with rainbow colors are the realizations.