Optimizing the Huff 'n' Puff Gas Injection Performance in Shale Reservoirs Considering the Uncertainty: A Duvernay Shale Example

Hamidreza Hamdi and **Christopher R. Clarkson,** University of Calgary; **Ali Esmail,** Ovintiv Corporation; and **Mario Costa Sousa,** University of Calgary

Summary

Recent studies have indicated that huff 'n' puff (HNP) gas injection has the potential to recover an additional 30 to 70% oil from multifractured horizontal wells in shale reservoirs. Nonetheless, this technique is very sensitive to production constraints and is impacted by uncertainty related to measurement quality (particularly frequency and resolution) and lack of constraining data. In this paper, a Bayesian workflow is provided to optimize the HNP process under uncertainty using a Duvernay shale well as an example.

Compositional simulations are conducted that incorporate a tuned pressure/volume/temperature (PVT) model and a set of measured cyclic injection/compaction pressure-sensitive permeability data. Markov-Chain Monte Carlo (MCMC) is used to estimate the posterior distributions of the model uncertain variables by matching the primary production data. The MCMC process is accelerated by using an accurate proxy model (kriging) that is updated using a highly adaptive sampling algorithm. Gaussian processes are then used to optimize the HNP control variables by maximizing the lower confidence interval (μ-σ) of cumulative oil production (after 10 years) across a fixed ensemble of uncertain variables sampled from posterior distributions.

The uncertain variable space includes several parameters representing reservoir and fracture properties. The posterior distributions for some parameters, such as primary fracture permeability and effective half-length, are narrower, whereas wider distributions are obtained for other parameters. The results indicate that the impact of uncertain variables on HNP performance is nonlinear. Some uncertain variables (such as molecular diffusion) that do not show strong sensitivity during the primary production strongly impact gas injection HNP performance. The results of optimization under uncertainty confirm that the lower confidence interval of cumulative oil production can be maximized by an injection time of approximately 1.5 months, a production time of approximately 2.5 months, and very short soaking times. In addition, a maximum injection rate and a flowing bottomhole pressure around the bubblepoint are required to ensure maximum incremental recovery. Analysis of the objective function surface highlights some other sets of production constraints with competitive results. Finally, the optimal set of production constraints, in combination with an ensemble of uncertain variables, results in a median HNP cumulative oil production that is 30% greater than that for primary production.

The application of a Bayesian framework for optimizing the HNP performance in a real shale reservoir is introduced for the first time. This work provides practical guidelines for the efficient application of advanced techniques for optimization under uncertainty, resulting in better decision making.

Introduction

One of the primary methods for maximizing well performance in low-permeability (tight and shale) oil reservoirs is multistage hydraulic fracturing in horizontal wells. Multistage fracturing can create a complex network of fractures because of the generation of new fractures and possibly reactivation of existing natural fractures. As a result, an increase in effective permeability in the near-wellbore area is observed. However, the ultimate oil recovery during primary production is still low using this method, and most in-situ oil remains in place in low-permeability reservoirs.

Several enhanced oil recovery (EOR) techniques, such as chemical methods and waterflooding, have been proposed to improve oil production from tight reservoirs. However, some practical issues such as capital and operational cost and conformance control have limited EOR application to mainly good-quality tight formations with larger pore throats (Alfarge et al. 2017b). Recently, the industry has been motivated to pursue miscible gas-injection EOR, particularly after EOG Resources reported successful pilot implementation of a cyclic gas-injection technique, with 30 to 70% increase in oil production (Rassenfoss 2017). Although, the details of EOG's pilot operations have not been made available to the public, analysis of production data (Hoffman and Evans 2016; Hoffman 2018) has suggested that the EOG's piloting method closely resembles a multiwell HNP gas-injection process.

A typical HNP process starts with injection of gas into the reservoir through the hydraulic fractures during the huff period. The well is then shut in for some time to allow the injected gas to soak and mix with the in-situ oil. Finally, the oil and gas are produced simultaneously during the production or puff period. This process is usually repeated several times until the incremental oil recovery after additional HNP cycles becomes negligible.

Although operational details for historically implemented HNP field cases are limited, simulation work and laboratory experiments have suggested that HNP is an effective approach to enhance the oil production from tight and shale reservoirs (Alfarge et al. 2017b). Important goals of simulation and experimental work are to evaluate the recovery mechanisms and explore the impact of various solvents and operational conditions on the oil production. For example, Song and Yang (2013) experimentally studied the impact of miscible and near-miscible carbon dioxide (CO_2) EOR, with continuous and cyclic gas-injection processes, for the Bakken core plug samples of southern Saskatchewan (Canada). Gamadi et al. (2013) used nitrogen in their HNP experiments for selected core plug samples from the Barnett, Marcos, and Eagle Ford shales and argued that higher pressures (and production drawdown), combined with longer soaking

This paper (SPE 195438) was accepted for presentation at the SPE Europec featured at 81st EAGE Conference and Exhibition, London, England, UK, 3–6 June 2019, and revised for publication. Original manuscript received for review 7 April 2020. Revised manuscript received for review 15 May 2020. Paper peer approved 16 June 2020.

times, could lead to an increase in production in the first few cycles. Later, Gamadi et al. (2014) used CO_2 as an alternative solvent in their HNP experiments and concluded that using higher pressures close to miscibility pressure, and longer soaking times, could enhance production in the first few cycles. They also observed that, for a fixed total operating time of HNP experiments, using a higher number of cycles with shorter soaking times could outperform HNP with fewer cycles and longer soaking times. In another study, Li and Sheng (2016) studied the effect of core plug size on the HNP oil recovery and concluded that a higher recovery can be achieved by using core plugs with larger surface-to-volume ratios and applying larger radial pressure gradients.

Although laboratory experiments, such as those described previously, can provide significant qualitative information about the effectiveness of the HNP process, as discussed by Li and Sheng (2016) and Alfarge et al. (2017b), there are significant challenges faced when scaling laboratory experiments, performed using small core plugs and under closed boundary conditions, to field scale. To assess the feasibility of HNP gas injection before field implementation, many authors (Wan 2013; Fai-Yengo et al. 2014; Pu and Hoffman 2014; Yu et al. 2014; Sheng 2017) have conducted numerical reservoir simulations to integrate various available data and to sensitize the performance of HNP using realistic operational scenarios. In particular, Yu et al. (2014) calibrated a numerical model using production data from the Middle Bakken Formation in North Dakota, which was then used to perform CO_2 HNP sensitivity analyses using ranges of model parameters and operational conditions. They concluded that oil recovery is sensitive to the number of cycles, injection rate and time, and diffusion coefficient of CO_2. Sheng (2017) used a history-matched model of the Bakken Formation (United States) and studied the impact of some operational conditions on the cyclic methane injection. He argued that molecular diffusion and soaking times have a minimal effect on HNP oil recovery. Alfarge et al. (2017a) also studied data from the Bakken Formation (United States) and used numerical simulations and experimental design to demonstrate that molecular diffusion has a significant impact on the performance of HNP gas injection processes. They argued that the molecular diffusion is scale dependent, and the results of the laboratory experiments cannot be directly applied to the field cases. A comprehensive review of the improved oil recovery techniques in tight and shale reservoirs in the North America is given in Alfarge et al. (2017b).

A survey of the literature reveals that there is no unique set of parameters and operational conditions that can be designed a priori to guarantee an optimal HNP performance in different field cases. To the best of our knowledge, despite numerous publications addressing the sensitivity of HNP performance with respect to some design parameters, there is no study that can systematically quantify the impact of uncertainty on designing an optimal cyclic gas-injection process. Accurate assessment of the feasibility and design of cyclic gas-injection processes in tight and shale reservoirs is a challenging task and is a function of various parameters, such as fluid interaction and reservoir characteristics. A rigorous data integration workflow should be used to assess the impact of uncertainty on cyclic gas-injection performance predictions. In our previous study (Hamdi et al. 2018a), a *deterministic* workflow was presented to evaluate the feasibility of HNP piloting using data from a multifractured horizontal well (MFHW) completed in the Duvernay shale in Canada. Various rock and fluid properties were measured in the laboratory and used as input to a numerical model calibrated using available (primary) production data. The calibrated numerical model was then used to study the impact of different solvents (lean and rich gases) on the cyclic gas-injection process. In this work, our previous study is extended by presenting a *stochastic* Bayesian framework to obtain a range of compositional history-matched models sampled from the posterior probability distributions. These models are used to correctly propagate the uncertainty and optimize the HNP performance under uncertainty. Such analysis can provide significant information on risks associated with implementing an HNP pilot that can lead to better decision-making practices. All the compositional simulations performed in this work use the tNavigator® (Rock Flow Dyamics, LLC, Moscow, Russia; Rock Flow Dynamics 2018) software, which is a multiplatform, multiformat, closed-loop reservoir-management package with a super-scalable parallel numerical engine.

Well and Reservoir Overview

The Duvernay shale is one of the most economically significant unconventional reservoirs in the Western Canadian Sedimentary Basin (Lyster et al. 2017). The Duvernay has extremely low matrix permeabilities (generally < 10 nd, as determined from laboratory analysis using crushed samples) and porosities (average of approximately 6%) (Fothergill et al. 2014). In-situ hydrocarbon fluid compositions range from dry gas to black oil. In the studied area, the average reservoir temperature and pressure are around 230°F and 8,700 psia, respectively.

The studied well is a 33-stage MFHW, with a lateral length of approximately 7,470 ft, which has been completed in the volatile oil window of the Duvernay shale. This well has produced oil, water, and gas for a period of 2.5 years under the primary recovery (depletion). **Fig. 1** provides the raw (unsmoothed) and smoothed flowing bottomhole pressures and production rates for the studied well. A filtering technique (moving average) was used to smooth the production/flowing pressure data for observing the trends (Hamdi et al. 2018a).

Theory

The workflow to optimize HNP performance under uncertainty requires several steps including a sensitivity analysis, Bayesian history-matching, and optimization under uncertainty (OUU). In the following, the theoretical background for implementation of each step of the workflow is briefly discussed.

Sensitivity Analysis. Sensitivity analysis methods are used to determine the impact of the model parameters (variables) on the output of a function (e.g., numerical simulator). The purpose of such an analysis is to understand which parameters have the greatest effect on the behavior of the underlying function. As an outcome, the behavior of a model can be described with only a few controlling variables (most impactful parameters). Fewer model parameters can in turn decrease the computational cost of the history-matching process by decreasing the number of simulations required.

There are several methods to perform sensitivity analysis. Some methods, such as Sobol' indices (Sobol' 1967), attempt to account for the interactions between the parameters and identify the nonlinear relationships between model input and output. However, this comes at the expense of conducting thousands of flow simulations. One way to decrease the computational cost associated with the sensitivity analysis is to use suitable proxy (surrogate) models such as polynomial chaos expansion (Blatman 2009; Christie and Bazargan 2012). However, because at this stage the number of model parameters is still large, we may need to perform many simulations before being able to construct an accurate proxy.

Another class of sensitivity analysis techniques is one-at-a-time sensitivity analysis methods. These methods are favorable because generally only a limited number of simulations are required to conduct the sensitivity analysis. However, these methods assume a linear relationship between model input/output and cannot quantify the complex (higher-order) interactions between the model parameters. Therefore, great care should be taken to ensure the sensitivity analysis does not lead to misleading results. Nonetheless, considering the pros and cons of different approaches, in this study, we obtain the initial sensitivity indices $S(i)$ using a one-at-a-time sensitivity analysis

technique known as the Cotter method (Cotter 1979). In this technique, we only need to perform $2M + 2$ simulations, with M being the number of parameters. The sensitivity indices are obtained as follows (Marelli and Sudret 2014; Lataniotis et al. 2015):

- Calculate the output of one run when all the variables are set to their lower values (y_{low})
- Calculate the outputs for M runs when all the variables are set to their lower values except for the subject parameter i, which is set to its higher value ($y_{\text{low}\rightarrow i\ \text{up}}$ and $i = 1,\ldots,M$)
- Calculate the outputs for M runs when all the variables are set to their higher values except for the subject parameter i, which is set to its lower value ($y_{\text{up}\rightarrow i\ \text{low}}$ and $i = 1,\ldots,M$)
- Calculate the output of one run when all the variables are set to their upper values (y_{up})
- Calculate the odd $[C_o(i)]$ and even $[C_e(i)]$ order effects and rank the parameters using the summation of these two effects.

$$S(i) = |C_o(i)| + |C_e(i)|, \quad \ldots\ldots (1)$$

where

$$C_o(i) = \frac{1}{4}[(y_{\text{up}} - y_{\text{up}\rightarrow i\,\text{low}}) - (y_{\text{low}\rightarrow i\,\text{up}} - y_{\text{low}})], \quad \ldots\ldots (2)$$

and

$$C_e(i) = \frac{1}{4}[(y_{\text{up}} - y_{\text{up}\rightarrow i\,\text{low}}) + (y_{\text{low}\rightarrow i\,\text{up}} - y_{\text{low}})]. \quad \ldots\ldots (3)$$

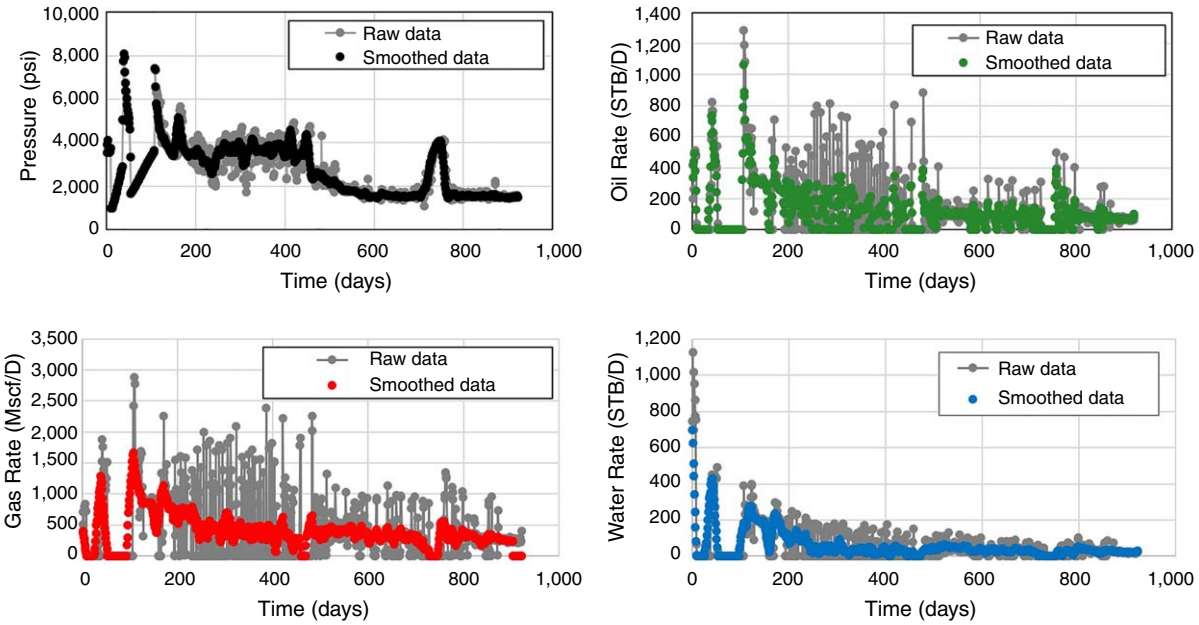

Fig. 1—Production data for the studied well. The raw data are shown using gray markers and lines and the smoothed data are represented with colored markers. From Hamdi et al. (2018a).

In these equations, $C_o(i)$ and $C_e(i)$ represent the expectation of the odd and even order effects involving parameter i, respectively. The sensitivity indices can be normalized using the summation over all the indices; that is, $S(i)/\sum S(i)$. Additional details and the derivation are given in Cotter (1979).

Bayesian History Matching. With traditional history-matching methods, we are mainly concerned with finding the best set of model parameters that can describe the observed production data. However, this approach will not allow us to correctly quantify the impact of uncertainty. This pointwise parameter estimation may not be very informative in many situations because, in practice, there might be a range of distinct (or combined) parameters that provide a good match to the data (Hamdi et al. 2017b). The implication of relying on only one best case is underestimation of the uncertainty resulting in poor model predictions. For example, some parameters may not be very sensitive to primary production data, but they might have a substantial impact on cyclic gas injection. Therefore, selecting arbitrary default values for insensitive parameters could have a large impact on model predictions for different production scenarios. In addition, performing sensitivity analysis to predict the future response using only prior ranges of the parameters undervalues the role of data to constrain the parameters. Furthermore, this simplified approach ignores the existence of nonlinearities that can frequently lead to over- or underestimation of the uncertainty in the prediction.

Bayesian inference is a statistically consistent approach that systematically quantifies the value of information, allowing knowledge about the model parameters ($\boldsymbol{m}$) to be updated by considering the existing data (D). The Bayesian framework is based on Bayes rule (Kruschke 2014) to estimate the posterior distributions using the likelihood (or misfit) function and our prior knowledge of the parameters. This is formulated as follows:

$$\Pr(\boldsymbol{m}|D) = \frac{\mathrm{L}(m;D) \times \Pr(\boldsymbol{m})}{\Pr(D)}, \quad \ldots\ldots (4)$$

where $\Pr(\boldsymbol{m})$ and $\Pr(\boldsymbol{m}|D)$ are the prior and posterior distributions, $\mathrm{L}(\boldsymbol{m}; D)$ is the likelihood, and $\Pr(D)$ is the evidence.

Prior Distribution. The prior distribution reflects our understanding of the model parameters without observing any data. When we have an extensive dataset, the impact of choosing the prior distribution on the shape of the posterior distribution is minimal. In contrast, when we have only collected sparse data, our initial belief can transfer substantial influence on the posteriors (Downey 2013). Therefore, it is important to properly formulate the prior distributions. However, in many situations we do not have any prior knowledge about the model parameters except a feasible initial range. For these cases, we can use uniform distributions to give equal weight to all possible values in the selected ranges of the model parameters. In other words, "we let the data speak for itself," and as a result, the shape of the posterior distribution will be determined by the likelihood function (Srivastava et al. 2014). Therefore, the mode of the posterior distribution, known as a maximum a posteriori estimate, will be the same as maximum likelihood estimation (Murphy 2012).

Likelihood Function. The likelihood function represents the probability of generating a realization of the data using the selected model when we use any instance of the model parameters. This is closely related to the history-matching misfit or objective function (*OF*). When history matching, it is assumed that the true reservoir response (*D*) can be approximated by the simulator's output (*Y*); that is, $D = Y + E$, in which *E* represents the measurement error and the simulation inaccuracy. If simulation error is neglected (e.g., because of gridblock size) and it can be assumed that the observed data set (i.e., production rates and measured pressure data) are mutually independent, the likelihood function can be represented by a multivariate Gaussian distribution (Bard 1974) as follows:

$$L(\boldsymbol{m}; D) = \prod_{\mathbf{j}=1}^{\mathbf{m}} \frac{(2\pi)^{-n_j/2}}{|C_{Dj}|^{1/2}} \mathrm{Exp}\left[-\frac{1}{2}(D_j - Y_j)^{\mathrm{T}} \boldsymbol{C}_{Dj}{}^{-1}(D_j - Y_j)\right]. \quad (5)$$

If it is further assumed that the measurement errors for all data points in any data set *j* are independent and normally distributed, that is, $E_j \sim N(0, \boldsymbol{C}_{Dj})$, the negative log-likelihood (or the regular misfit function) can be simplified as follows:

$$-\mathrm{Ln}(L) \propto \frac{1}{2}\sum_{j=1}^{m}\sum_{i=1}^{n_j}\left[\frac{(d)_{i,j} - (y)_{i,j}}{\sigma_{i,j}}\right]^2, \quad (6)$$

where σ is the measurement error for any data point, *m* is the number of data sets (such as pressure and production rates), *n* is the number of timesteps, and *d* and *y* are the measured and simulated data points, respectively. In the history-matching framework used herein, the following objective function (or total misfit) is applied:

$$OF = \frac{1}{2}\sum_{i=1}^{n}\left(\frac{q_{oi,m} - q_{oi,\mathrm{sim}}}{\sigma_{oi}}\right)^2 + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{q_{gi,m} - q_{gi,\mathrm{sim}}}{\sigma_{gi}}\right)^2 + \frac{1}{2}\sum_{i=1}^{n}\left(\frac{q_{wi,m} - q_{wi,\mathrm{sim}}}{\sigma_{wi}}\right)^2, \quad (7)$$

where q_o, q_g, and q_g represent oil, gas, and water rates, respectively. The subscripts *m* and sim indicate measured and simulated data, respectively. The measurement errors (σ) are estimated using a moving average (or median) filter as suggested by Aanonsen et al. (2003). The objective function is evaluated using the unsmoothed observations.

MCMC. A major issue with applying Bayes theorem to multidimensional problems is the difficulty in analytically calculating the evidence $\Pr(D) = \int \mathrm{L}(\boldsymbol{m};D) \times \Pr(\boldsymbol{m})\mathrm{d}(\boldsymbol{m})$ because it involves the computation of a complex multiple integral. However, statistical analysis shows that making inference about model parameters does not really require the calculation of evidence. In statistical inference, the evidence is seen as a normalization constant to ensure the resulting posterior distribution becomes a true probability density (i.e., it should integrate to 1). However, numerical evaluation of the posteriors requires a large number of samples from $\mathrm{L}(\boldsymbol{m};D) \times \Pr(\boldsymbol{m})$. MCMC methods (Gamerman 1997) is a class of Monte Carlo sampling approaches that can generate a large chain of weakly dependent samples from the parameter space to converge to a stable target (posterior) distribution (Privault 2013). Several efficient MCMC methods have been developed in the literature to efficiently sample from the posterior distributions. A comprehensive review of different MCMC methods with application to petroleum engineering problems is presented in Goodwin (2015). In this study, a multichain Gibbs sampler (Geman and Geman 1984) is used to estimate the posterior probability distributions.

Global Proxy. Application of MCMC requires extensive sequential sampling from $\mathrm{L}(\boldsymbol{m};D) \times \Pr(\boldsymbol{m})$ to converge. This means that a large number of numerical simulations need to be conducted to evaluate the likelihood at various suggested points, which is not practically feasible. To alleviate the computational cost associated with numerical Bayesian inference, a global surrogate (or proxy) is constructed and used as an alternative to conduct extensive numerical simulations. A global proxy is a cheap-to-evaluate function that can mimic the behavior of a numerical simulator in the entire domain (Crombecq et al. 2009). Examples of surrogate models to represent the misfit function are kriging (Chiles and Delfiner 2012) and radial basis functions (Buhmann 2003).

One way to construct the global surrogate model is to use some one-time space filling designs such as Latin hypercube (Ye 1998) to generate a large number of samples a priori and run the numerical simulation on all the samples to evaluate the misfit values. However, this method not only is computationally expensive but also does not always correctly capture the nonlinearities in the domain.

Other approaches are to include the proxy model within the history-matching process and use the proxy to generate different history-matching models. In these methods, the proxy model is improved at different iterations of the history-matching workflow. A survey of the literature shows that many published works discuss the benefits of using proxy models for Bayesian history matching and uncertainty quantification. An excellent example of using the proxy models for facilitating the optimization and uncertainty quantification in reservoir engineering is given by Slotte and Smorgrav (2008). They suggest that the combined use of the kriging model and the MCMC approach during history matching can decrease the computational cost of the uncertainty quantification significantly. The proxy acceptance-rejection method is also proposed by Yang et al. (2015), which uses the radial-basis function neural networks to construct an adaptable proxy and generate multiple history-matching methods for an Eagle Ford Shale case. Schulze-Riegert et al. (2016) provided an excellent example of combining the proxy model and MCMC for sampling and generating multiple history matching in a North Sea example. Wantawin et al. (2017a, 2017b) used a design of experiment approach with a quadratic and high-order polynomial response surface to obtain multiple history-matching models for some MFHWs in the Marcellus and Bakken reservoirs. An excellent overview and performance comparison of various proxy-based uncertainty quantification techniques is given by Goodwin (2015). In all these methods, the proxy model accuracy is improved in each iteration and is used to generate multiple calibrated models, which can be used to predict future responses probabilistically. One drawback of these methods is that unless lots of simulations are sampled, the final history-matched models may not always capture the full posterior distributions, and because they target to decrease the misfit functions, the parameter space may not be explored well.

In this work, an adaptive sampling method is used instead to put more samples in the areas that are more important and/or undersampled. The kriging surrogate model is used, and the LOcal Linear Approximations (LOLA)-Voronoi (LOLA-Voronoi) sampling

method (Crombecq et al. 2009; Gorissen et al. 2010) is employed to provide a balance between exploration and exploitation in the sampling. The algorithm starts with generating only a few samples from either a space-filling design and/or the results of a population-based history-matching technique such as differential evolution (Storn and Price 1995; Hamdi et al. 2015). More samples are sequentially added based on the LOLA-Voronoi algorithm. In the LOLA part of this algorithm **(Fig. 2, left),** a local linear adaptive method is used to linearly estimate the gradient at a sample point from the nearby samples using the least-squares method. The estimated gradient is then used to linearly estimate the function values in nearby samples for which the actual values are known. If the estimates convey a large error in some areas around the subject sample point, it indicates a potential nonlinearity in that direction, and subsequently more samples are suggested to be placed around that subject point. The Voronoi part of the algorithm (Fig. 2, right) discretizes the domain around the sample points using Voronoi tessellation. If the volumes of some Voronoi cells are larger than others, it indicates potential undersampled locations, which would be good candidates for placing additional samples in those areas. These two elements of the algorithm work together and balance between local refinement or global exploration of the domain. By repeating this algorithm with an affordable number of iterations, the kriging surrogate model is dynamically tuned to better represent the numerical simulator.

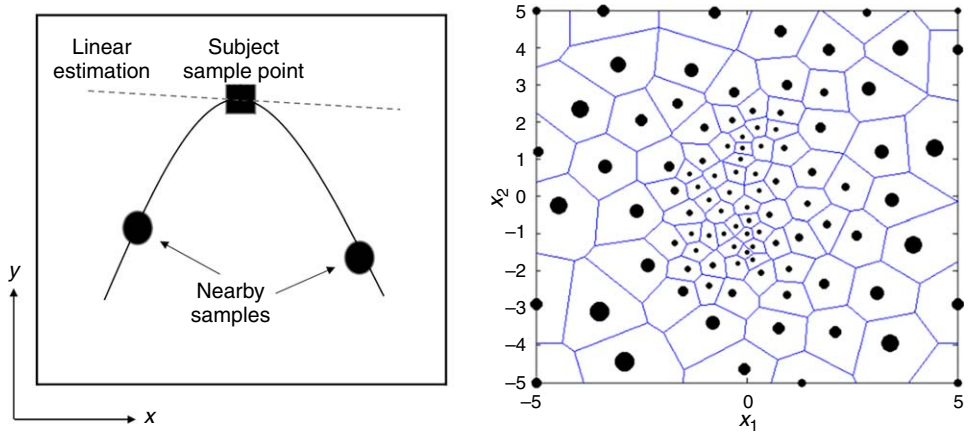

Fig. 2—LOLA-Voronoi adaptive sampling method. In the LOLA element of this algorithm (left) the local nonlinear areas near the subject sample points are identified. In the Voronoi part (right), the Voronoi cells are generated around all samples. The samples with larger volumes are shown with larger circles. In each iteration, a tradeoff between these two elements is sought, and a new sample point is suggested. The numerical simulator is evaluated at this point and the point is added to the pool of data. Images modified after Crombecq (2011).

As stated earlier, there are many different methods in the literature that are introduced to help the Bayesian inference and probabilistic forecasting using the proxy models. Yet, it is crucial to notice that according to the "no free lunch" theorem in optimization (Wolpert and Macready 1997), there is no single approach that can be applied to all the problems successfully. Each method has its own set of pros and cons and may suit some specific problems better than the others. However, this is not the objective of this paper to compare various proxy and sampling strategies. The LOLA-Voronoi method has only been chosen because it has some interesting properties that were found to suit the objective of this specific study better. In the LOLA-Voronoi method, the aim is to create a global proxy that can replace the simulator independent of the status of history matching. The automatic exploration and exploitation components of the algorithm are designed to capture possible multimodalities in the likelihood function. However, if for some specific problems, the objective was to only focus near the maximum likelihood areas, the LOLA component can be given more weight in the process. Providing additional samples near the maximum likelihood areas, such as those obtained from the previous optimization methods, can also be useful.

When the LOLA-Voronoi method is used, the inference is conducted using MCMC separately. Hence, the used proxy model in this study is different from other popular methods (e.g., proxy acceptance-rejection) in the way that the proxy is constructed and used. In those methods, the proxy is involved within the history-matching process and is combined with some sampling methods to draw some samples from the estimated posteriors to provide various history-matching models. However, there might be some issues with this mechanism. For example, in the proxy acceptance-rejection method, there is a chance that the initial Latin hypercube design and/or the constructed proxy method may not fully represent the nonlinearities of the problem. In addition, the initial samples may only capture the areas that cannot fully cover the posterior regions (Mariethoz and Caers 2014). Therefore, it becomes very difficult for the process to progress toward better sample points in an affordable number of iterations.

OUU. The estimated posterior distributions can be used to provide a probabilistic forecast by providing the full probability distribution of the prediction or even simpler statistics such as P10, P50, and P90. This means that some of the posterior (equi-probable) models perform better than the others, and therefore, the decision makers are able to include the risk associated with their selection of the models that are used for future development strategies. OUU is a probabilistic approach (Bailey et al. 2011) that is used to optimize the control variables (CVs) when there exists a set of uncertain variables (UVs).

OUU has been applied to many conventional and several unconventional reservoir problems to assist the history-match process and/or to optimize the CVs while there is a persistent uncertainty in the parameters. The OUU framework has been frequently applied to well-placement optimization problems under geological uncertainty (Badru and Kabir 2003; Morales et al. 2011). Chen et al. (2009) applied an ensemble-based optimization (EnOpt) method along with ensemble Kalman filter for a closed-loop optimization problem to optimize the geological distribution and the production constraints simultaneously. In a similar type of problem, Yeh et al. (2014) proposed a dynamic fingerprinting approach that combines both MCMC and the experimental design approach to history match the production data considering geological heterogeneity. Podhoretz and Valkó (2014) applied an OUU framework for decreasing the risk of a massive fracturing design in an offshore case in the Gulf of Mexico. The application of OUU methods has also been extended to predict some EOR mechanisms in conventional and heavy oil reservoirs (Chugunov et al. 2015; Ampomah et al. 2016).

The OUU framework has also been implemented for several problems in unconventional reservoirs. Tilke et al. (2015) applied an OUU routine for an automated field development planning project using the decline curve techniques. Nguyen et al. (2016) implemented a robust optimization approach to include the impact of geological uncertainty on optimizing the hydraulic fracturing and continuous gas-injection design for a shale reservoir model. Mehranfar et al. (2018) used an OUU scheme to optimize the well and fracture cluster spacings under parametric uncertainty to minimize the development cost. In all these examples, the primary concern of using an OUU framework is to decrease the risk and avoid the failure of a future project by including uncertainty in the decision-making process.

The objective of this study is to design an HNP process with a set of optimal operational conditions that can maximize the cumulative production after 10 years of operations. One way to do this is to pick the best history-match model with the lowest misfit and use it to find a set of parameters that can give the highest cumulative oil production (Hamdi et al. 2018a). However, this strategy excludes the impact of uncertainty on the optimization process.

The terms CV and UV are used to distinguish between the operational parameters that can be controlled and variables for which the uncertainty cannot be decreased. For the HNP optimization problem, the CVs include the injection rate and injection time for the huff period, soaking time, and bottomhole production pressure during the puff period. The uncertain variables include the posterior probability distributions obtained from applying the MCMC and the other uncertain parameters that did not show any large impact during the history-matching process.

Using the OUU scheme, the lower confidence interval of the cumulative oil production after 10 years is optimized. This is represented by $F = \mu - \sigma$, where μ and σ are the mean and standard deviation of the cumulative oil production obtained from any set of CVs with a fixed population of samples from UVs, respectively. In other words, the intent is to find a set of CVs that, in under some worst-case scenarios (pessimistic combination of possible UVs), can give the best performance. The lower confidence interval is schematically shown in **Fig. 3.**

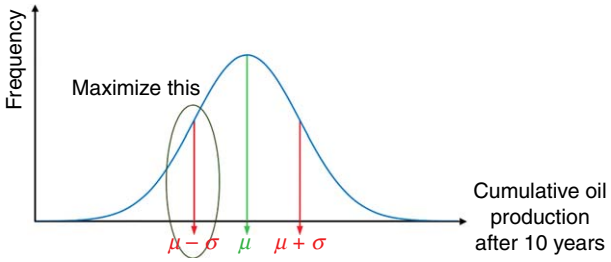

Fig. 3—Schematic representation of the lower confidence interval that is used in the optimization process.

For each iteration of the optimization algorithm, a new set of CVs is proposed. This set of CVs is combined with many samples from UVs that are invariant over the entire optimization process. The purpose of this pool of UVs is to provide a reasonable statistical representation to cumulative oil production for any set of CVs.

In theory, a large number of samples from the UV space should be selected to be able to infer accurate statistical parameters (i.e., σ and μ). However, this makes the optimization routine largely impractical. For example, if 500 samples are selected from UVs and optimization is performed for at least 200 iterations, approximately 100,000 numerical simulations should be run (which is computationally expensive). Therefore, either the dimension of the problem should be decreased (e.g., by only considering the most uncertain variables) and/or efficient one-time sampling algorithms should be used to adequately approximate the statistics.

Gaussian Process-Based Optimization Method. It is worth noting that selection of an efficient optimization algorithm, which requires fewer iterations (simulations), can substantially decrease the computation cost of OUU. Robust population-based algorithms, such as differential evolution, usually require many iterations, whereas the popular gradient-based techniques cannot easily converge to *global* optima. In this paper, an efficient surrogate-based optimization method that uses Gaussian processes (GPs) (Hamdi et al. 2017a, 2017b, 2018b) is used to perform the optimization task. This algorithm is extremely efficient for optimizing computationally expensive problems with low to medium dimensions.

The benefit of using a GP-based history-matching approach over other proxy-based methods is two-fold: to conduct the expensive history-matching tasks using a fewer number of simulations and to sequentially construct and validate a global proxy model. Starting from only a few simulations, the GP model is trained to guess the next best sample location. A key difference between the GP-based approach and any other proxy-based optimization method is that not only can the currently tuned GP model deliver a prediction for an unknown location, but it also can provide a measure of uncertainty for that prediction. These two aspects of the GP model (i.e., prediction and uncertainty) will be used within an auxiliary statistical infill criterion to propose the next best possible location to sample and sequentially retune and improve the predictability of the GP model. The actual simulation is then conducted on this suggested sample, and the initial data set is augmented by adding this new point. This process can be repeated over and over. This proxy-based optimization method is very efficient in finding global optima in fewer iterations.

The performance of the GP-based optimization method in finding the global minimum of a one-dimensional function is schematically shown in **Fig. 4.** In this simple example, the physical model (i.e., the truth objective function) is shown by a solid black curve. The objective is to find the global minimum of this function with only a few iterations. Solid black dots indicate the currently available samples. At iteration i, a GP model is fitted to the data (the dotted red curve). This curve is used to predict the value of an unknown function (y) at any location (x). As shown in Fig. 4 (left), if only the fitted curve is used in the optimization, the global minimum, which is located in an unsampled area, can never be found. In other words, we cannot only rely on the fitted proxy to select a best sample location. The selection mechanism should be able to also account for uncertainty in the fitted proxy model. Therefore, in the GP-based optimization methods, some auxiliary functions with special space filling characteristics are introduced to help find the next best samples by including the uncertainty of the proxy.

Fortunately, the GP model can provide a prediction and a measure of uncertainty using a multivariate Gaussian distribution function $N[y^*, s^2(x^*)]$ for any location of interest (x^*). The extent of uncertainty in predicting the function value at point x^* is shown using a gray Gaussian curve (Fig. 4). This uncertainty measure is used to calculate the multimodal expected improvement (EI) (Jones et al. 1998; Jones 2001) function for any location in the parameter space (the blue curve). The maxima of the EI are located in the areas close

to the local minima or the unsampled areas with large standard deviations $s^2(x^*)$. As shown in Fig. 4 (left), in the first iteration, maximizing the EI has led to adding a point (x^{**}) close to the current best function value (y_b) (i.e., exploitation has overcome the exploration). This new sample point is evaluated using the truth function and is added to already available samples, which are then used to update the GP proxy. In the next iteration $i+1$, the EI has a higher maximum at x^{**}, which is identified with a large standard error (Fig. 4, right). In this iteration, the exploration has overcome the exploitation. At this stage, the convergence has achieved, and no further iteration is required.

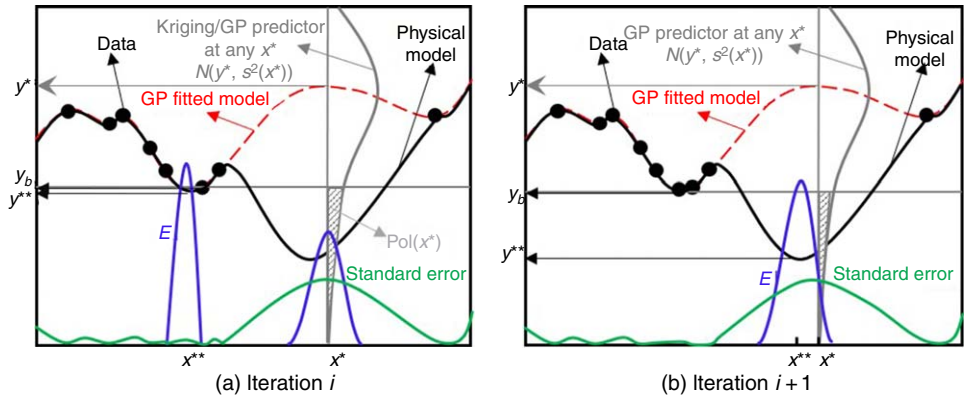

Fig. 4—Schematic representation of the GP-based optimization method in iterations *i* (left) and *i* + 1 (right).

There are many statistical criteria such as EI, probability of improvement (Jones et al. 1998; Jones 2001), upper confidence bound (Srinivas et al. 2010), and so on. These infill criteria can help the performance of the optimization process to find the global optima. EI can frequently outperform the probability of improvement as the latter can have some issues in exploring the space (Jones 2001). Upper confidence bound criterion is a robust infill criterion, and in a number of applications, its performance is comparable to the EI (Srinivas et al. 2010). In this study, the EI is used as it is easy to implement and has a favorable characteristic to balance the exploration and exploitation of the space.

The smoothness of the samples drawn from a GP model is controlled by some kernel functions. These kernel functions are similar to the variograms used in geostatistics. Although in three-dimensional geostatistical applications, the type of kernel and its hyperparameters (e.g., correlation lengths) are determined by fitting a variogram to some spatially observed samples, in the optimization process, that information is not readily available. Fortunately, many published works have demonstrated that the use of Gaussian or Matérn kernel functions (Rasmussen and Williams 2005) provides favorable results (Wilson and Adams 2013). Matérn kernel functions are usually preferred over Gaussian kernels for many practical applications (Stein 1999; Snoek et al. 2012) because they do not imply any infinite differentiable characteristics (Mohammad and Nishida 2016). The hyperparameters are optimized using a maximum likelihood estimation procedure every time the GP proxy is constructed.

The preference to choose the GP-based optimization over any other popular gradient-based or population-based method is highlighted in expensive global optimization problems. OUU is a perfect example problem because we cannot afford to run a large number of expensive compositional simulations. The GP-based optimization method has the ability to balance between the exploration and exploitation aspect of the sampling and identify the global optima fairly quickly. However, if there is no limitation in the number of simulations or if the objective function is convex, other optimization methods can be readily used instead of this proxy-based approach. Each technique comes with its own set of pros and cons, which should be considered before choosing any optimization method.

Additional details on the GP-based optimization method are given in Hamdi et al. (2017a, 2018b), where the superiority of this method compared to other optimization methods and its application to several benchmarking problems are also discussed.

Methods: Numerical Model Setup

In the following section, the fluid model construction and the numerical model setup are described to perform the various steps of the workflow used to maximize HNP oil recovery for the studied MFHW.

Fluid Model. The reservoir fluid corresponds to a volatile oil reservoir with an API oil gravity of approximately 49.1. The bubblepoint pressure of the fluid is 3,724 psia. The Peng-Robinson equation of state (Robinson and Peng 1978) is used to tune the fluid model using several available laboratory experiments including saturation pressure, constant composition expansion, constant volume depletion, and separator tests. The equation of state comprises seven pseudocomponents **(Table 1). Fig. 5** provides the pressure/temperature diagram (phase envelope) of the initial reservoir fluid.

Component	Molar Composition (%)
CO_2	0.43
N_2 to C_1	49.34
C_2 to nC_4	25.79
iC_5 to C_7	8.03
C_8 to C_{12}	9.51
C_{13} to C_{19}	3.93
C_{20+}	2.98

Table 1—Overall composition of the pseudocomponents in the in-situ fluid.

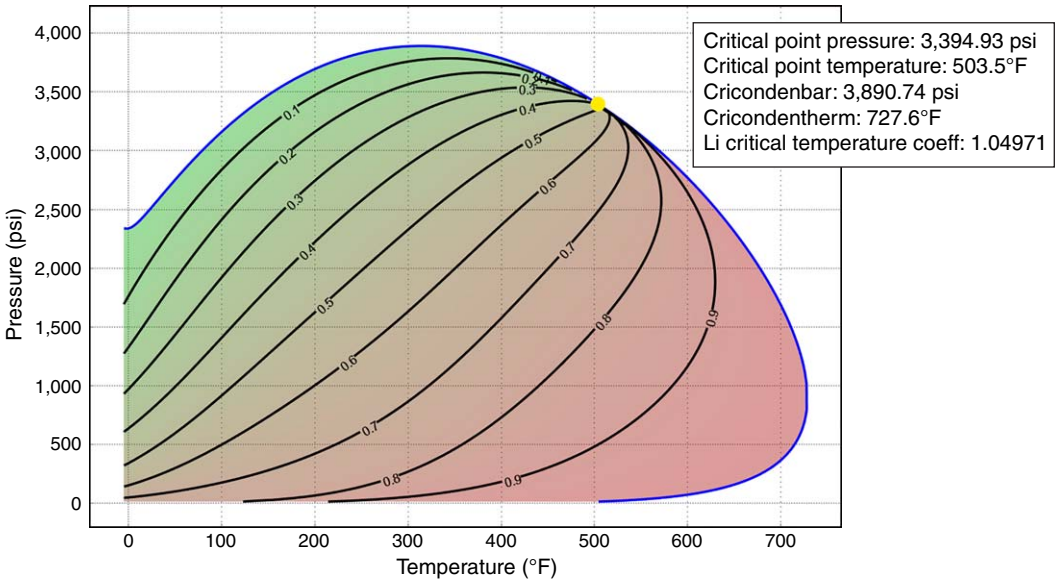

Fig. 5—Phase envelope for the in-situ reservoir fluid (volatile oil).

It is worth noting that, at the time of this study, no EOR-related PVT experiments (e.g., multicontact tests) were available. Therefore, the PVT model is not tuned to those experimental results.

Numerical Model. As discussed in earlier work (Hamdi et al. 2018a), a limited enhanced fracture region (LEFR) model (Hao et al. 2016, Hamdi et al. 2018a) was found to be suitable for history matching the subject well. This model assumes planar symmetric fractures with the possibility of having an enhanced fracture region with a limited extent along the primary hydraulic fracture. Therefore, only flow toward one side of a one-quarter fracture is modeled. By doing so, the well production rates are adjusted to represent a one-quarter fracture model with an assumed (typical) efficiency of 50% per fracture stage (estimated from production logging data). The latter assumption leads to 66 planar symmetric fractures, which are assumed to be uniformly distributed along the wellbore.

The base case model includes five primary reservoir layers with distinct petrophysical properties (effective porosity and permeability) obtained from laboratory and log correlations (see Table 1 in Hamdi et al. 2018a). The hydraulic fracture is modeled using a uniform slab with a width of 1 ft. The fracture half-length is fixed at 500 ft to honor microseismic interpretations and field observations. Logarithmic gridding is used in both x- and y-directions to accurately capture the flow toward the hydraulic fracture and the perforations.

Fig. 6 shows a typical LEFR model that is used for the simulations and the optimization algorithm. In this model, stress-dependent conductivities for the propped fracture and rock matrix are included, which are measured in the laboratory under various loading and unloading stress conditions. The stress-dependent curves are normalized to represent the transmissibility multipliers with hysteresis. The hysteresis effect reflects the irreversible paths caused by proppant embedment and crushing with stress loading/unloading that can happen in drawdown and build-up cycles during primary production or cyclic gas-injection processes. **Fig. 7** provides the stress-dependent transmissibility multipliers for the propped fractures and the shale matrix.

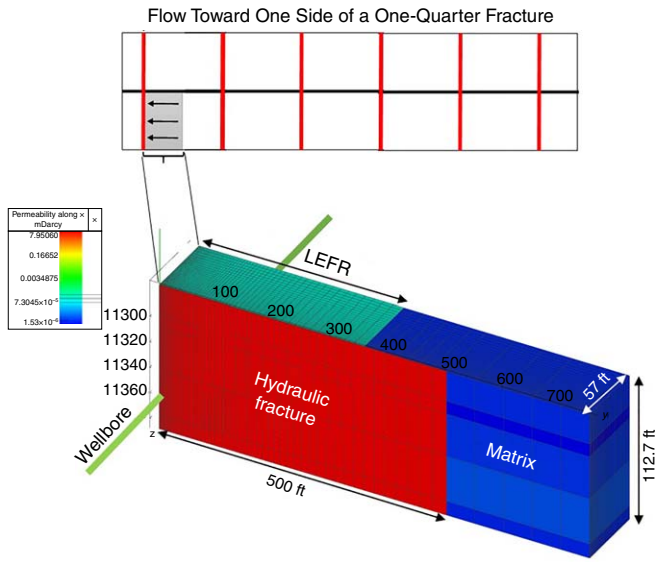

Fig. 6—MFHW with planar symmetric fractures (top), and an equivalent LEFR model (bottom) to model the flow toward one side of a one-quarter fracture model. The three-dimensional model includes the hydraulic fracture (red) with a half-length of 500 ft, an enhanced fractured region with limited extent for history-matching (green), and the rock matrix (blue) to allow the flow from the tip of the fracture.

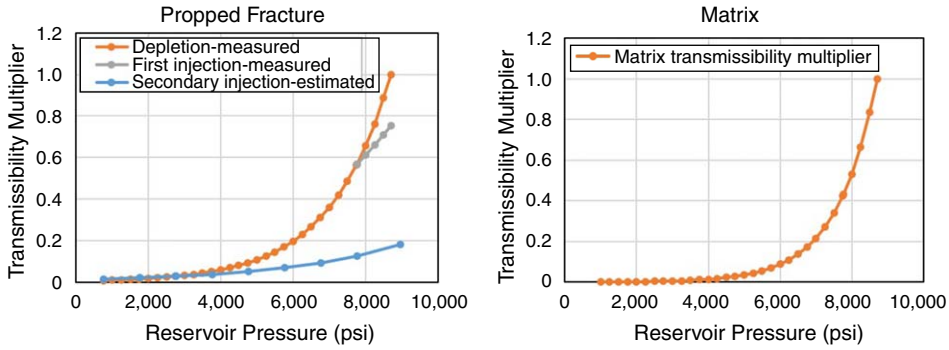

Fig. 7—Derived transmissibility multipliers for the propped fracture with different stress loading/unloading cycles (left) and the rock matrix with stress loading only (right). Transmissibility multipliers are obtained from the slopes of the measured stress-dependent permeability measurements in the laboratory (Hamdi et al. 2018a).

Results

Sensitivity Analysis. A set of uncertain variables are carefully selected to populate the LEFR model for numerical simulations and sensitivity analysis. These variables include the effective permeability of the propped fracture, an initial uniform water saturation in the propped fracture (S_{wHF}), the extent of the enhanced permeability region ($x_{f\text{-eff}}$), a permeability multiple to increase the effective permeability of this region, a similar molecular diffusion coefficient (D) for top five lighter gas components, and a set of parameters to generate the two-phase relative permeability curves for the matrix and the hydraulic fracture using the Corey-type equations (Christiansen 2007; Hamdi et al. 2018a).

The molecular diffusion coefficients for the components in the oil phase (D_o) are usually an order of magnitude smaller than those in the gas phase (Schlumberger 2011; Rock Flow Dynamics 2018). Therefore, in the simulations, it is assumed that $D_o = D/10$. In addition, as discussed in the previous study (Hamdi et al. 2018a), most of the produced water is assumed to be the flowback water that exists in the hydraulic fracture, and therefore the simulations are conducted in the absence of any mobile water in the shale matrix. Therefore, the oil-water relative permeability curves in the matrix have no influence on the simulation results. Moreover, the matrix porosity values are corrected to account for the decrease in pore volume caused by measured connate water saturation in the matrix. Hence, no (effective) connate water saturation presents in the simulations.

Table 2 lists the uncertainty range of all the variables used in the sensitivity analysis. The ranges of the variable are selected based on our general understanding of the problem to provide a relatively wide range for sensitivity analysis. In practice, additional available sources of information (e.g., analogue wells and laboratory measurements) can also be used to help constrain the variability of the parameters within their plausible ranges to avoid introducing unrealistic uncertainties.

Parameter	Minimum	Maximum	Parameter	Minimum	Maximum
n_{wF}	1	3	S_{rogM}	0.3	0.5
n_{owF}	1	2	S_{wHF}	0.4	1
S_{rowF}= S_{rogF}	0.05	0.25	$x_{f\text{-eff}}$ (ft)	100	400
S_{wcF}	0.1	0.3	Permeability multiplier or k_{mult}	10	60
k_{rwmaxF}	0.1	0.9	log(fracture permeability) or log(K_{HF})	−2	1.7
n_{ogF}	1	3	Gas diffusion coefficient D (cm^2/s)	10^{-4}	10^{-2}
n_{gF}	1	2	Default k_{romaxM}	1	1
S_{gcF}	0.01	0.1	Default S_{rowM}	0.5	0.5
k_{rgmaxF}	0.4	0.9	Default k_{rwmaxM}	0.5	0.5
n_{ogM}	3	7	Default S_{wcM}	0	0
n_{gM}	2	5	Default x_f (ft)	500	500
S_{gcM}	0.05	0.2	Default n_{owM}	5	5
k_{rgmaxM_mult}*	1	10	Default n_{wM}	3	3

M, matrix; F, fracture.

* $k_{rgmaxM} = k_{rgmaxF}/k_{rgmaxM\_mult}$.

Table 2—Range of input parameters used for sensitivity analysis.

The sensitivity analysis is conducted using the Cotter method (Cotter 1979) to evaluate the impact of model parameters on the total misfit (Eq. 7). **Fig. 8** shows the normalized sensitivity indices (Marelli and Sudret 2014). In this figure, the most impactful parameters are identified by red bars. These parameters include the effective permeability of the hydraulic fracture, the initial water saturation in the fracture, permeability (multiplier) of the enhanced fractured region, $x_{f\text{-eff}}$, and the oil-water relative permeability of the hydraulic fracture. This set of parameters will be used in the subsequent Bayesian history-matching analysis, whereas the rest of the (less impactful) parameters are fixed to their default values.

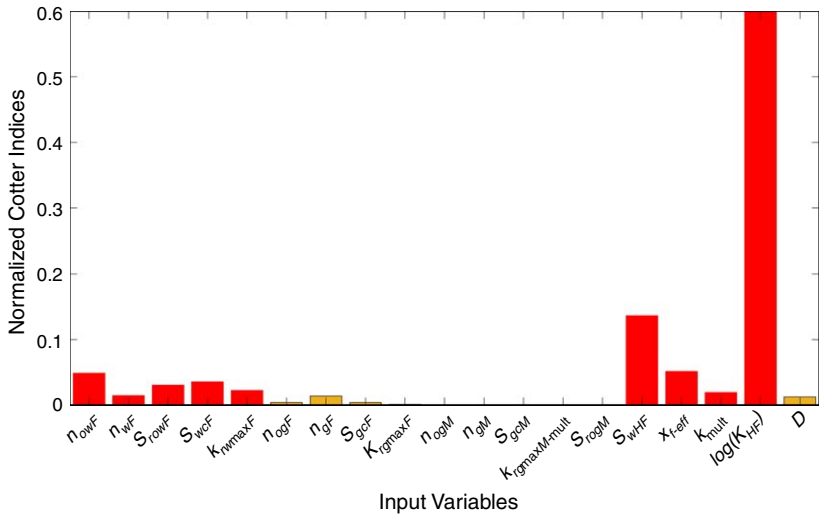

Fig. 8—Normalized Cotter sensitivity indices. The major impactful parameters are shown in red.

Bayesian History Matching. The more impactful parameters can be used to conduct numerical history matching and obtain the most likely parameters that can provide the best match to the data. With population-based methods such as differential evolution, and using uniform priors, one can sometimes obtain a set of solutions (models) sampled near the maximum a posteriori estimate. When using uniform priors, these sets of solutions are very helpful to generate an accurate proxy model to represent the likelihood function. Therefore, an initial history-matching exercise is performed using an adaptive differential evolution algorithm (Gong et al. 2014) with a maximum of 450 simulations. The samples are enriched with a two-level full factorial design to account for the combination of parameters with their values set at the highest and lowest bounds. This pool of samples is further augmented by adding the LOLA-Voronoi samples with 400 more samples. Eventually, a kriging model is fitted to all the samples to approximate the square root of the normalized misfit function. This proxy is used to define a log-likelihood function for use in the subsequent MCMC runs. The good quality of the kriging model in predicting the actual misfit is visually demonstrated in **Fig. 9** by displaying the cross-validation error histogram (left) and a prediction scatter plot with a diagonal trend (right). The figure illustrates that the errors are close to zero (Fig. 9, left) and that the proxy function can predict the majority of actual data points within one standard deviation of the kriging error (green whiskers in Fig. 9, right). Having a higher accuracy near the origin of the prediction scatter plot ensures the proxy can adequately represent the higher likelihood regions.

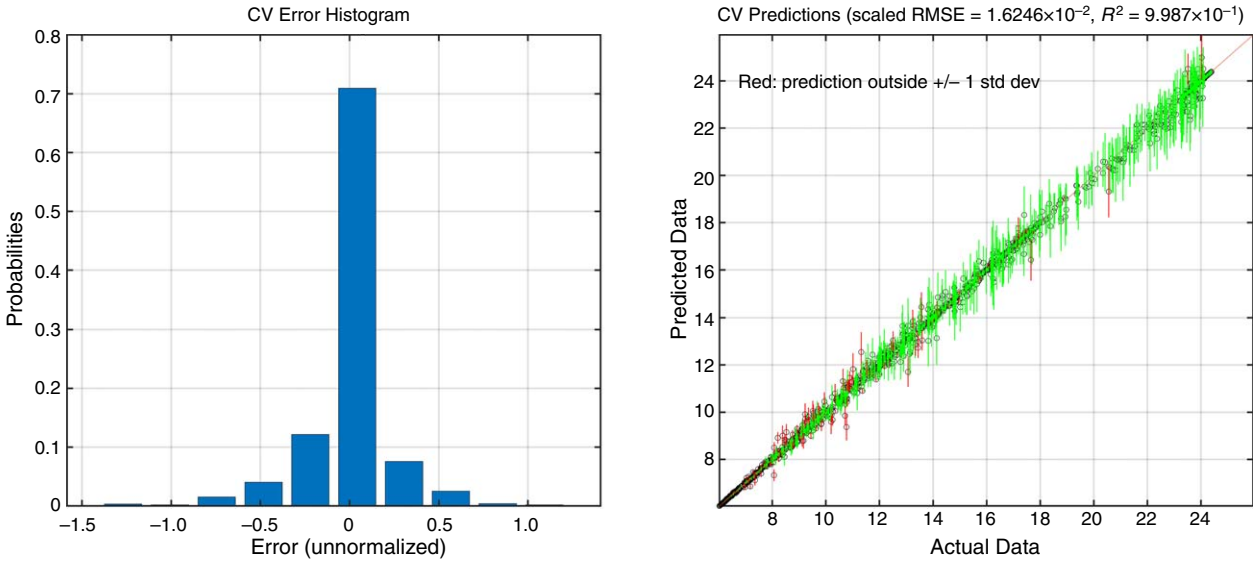

Fig. 9—The cross-validation error histogram (left) and the prediction error scatter plot (right). The green and red whiskers show the prediction within and outside one standard deviation (std dev) of the estimated values, respectively. RMSE = root mean square error.

Fig. 10 shows the projected mutual prior (the lower triangle) and the estimated posterior distributions (the upper triangle) obtained from a long 5-chain MCMC sampling using the Gibbs sampling method (Tong 2018). The figure highlights the value of production data in constraining the uncertainty in the parameter space. A narrower posterior portability distribution for some parameters such as permeability of the hydraulic fracture illustrates that the production data carries more information to constrain the permeability distribution. Any numerical model that is populated with the parameters sampled from the posterior distributions, can adequately match the measured data **(Fig. 11).** These posterior distributions, as well as uniform distributions for other less impactful uncertain variables, define an uncertain region that is used to optimize the HNP performance under uncertainty.

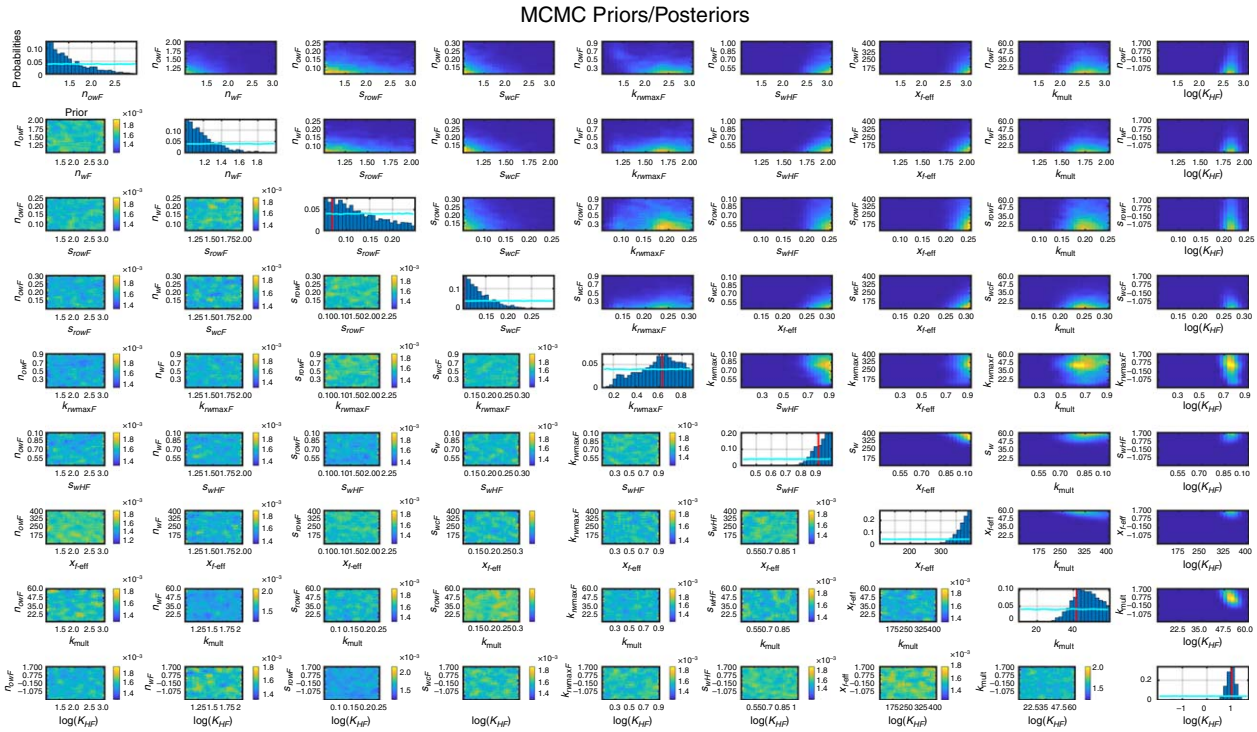

Fig. 10—Matrix of prior (lower triangle) and posterior probability distribution (upper triangle) projected in two dimensions.

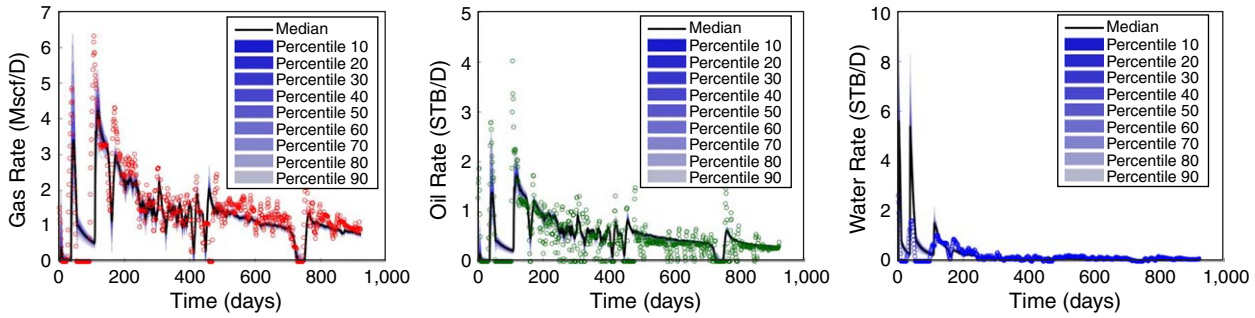

Fig. 11—Fan chart of probabilistic history-matching results using the posterior distributions. The production rates correspond to a one-quarter fracture model. The blue shaded regions are the percentiles around median.

HNP Optimization under Uncertainty. A quasirandom sequence (Saltelli et al. 2008) is used to generate a set of 50 uniformly distributed samples from UVs. The efficiency of this relatively small (but accurately designed) population of UVs in representing the uncertainty was checked against a larger population of 500 samples, yielding similar results. The UVs include the posterior distributions and the uniform ranges for other less (or un-) impactful parameters as shown in Fig. 8. Because of the high computation cost of performing OUU, optimization is performed for only the rich gas injection only. However, in the Discussion section, the performance of the HNP process using lean gas is commented on. The molar compositions of the lean and the rich gas used in the simulation are listed in **Table 3.**

Component	Lean Gas (Molar Composition, %)	Rich Gas (Molar Composition, %)
CO_2	0.718	0.682
N_2 to C_1	83.792	72.514
C_2 to nC_4	15.375	25.179
iC_5 to C_7	0.114	1.223
C_8 to C_{12}	0.001	0.402
C_{13} to C_{19}	0	0
C_{20+}	0	0

Table 3—Composition of lean and a rich gas used in the HNP simulations.

The optimization process tends to maximize the lower confidence interval of the cumulative oil production ($F = \mu - \sigma$) from the cyclic rich gas injection after 10 years of operation. This is estimated by calculating the mean and standard deviation of the cumulative oil production from an ensemble of 50 numerical models sampled from UVs with the same set of CVs. The GP-based optimization algorithm is applied, starting with an initial population of 20 CVs (i.e., repeated for an additional 120 iterations). This means that a total of 6,000 (50×120) simulations is performed during the course of the optimization process. In this problem, the GP model is constructed using a Matérn kernel with $\nu = 3/2$. This kernel has shown good performance and favorable convergences in many practical problems (Picheny et al. 2013; Marmin et al. 2015; Hamdi et al. 2017a). In addition, the EI statistical criterion is used to assist in the optimization process. Additional details on the effect of selecting different kernel function and using different statistical criteria on the performance of GP-based optimization method is discussed extensively in Rasmussen and Williams (2005) and Hamdi et al. (2017a). **Fig. 12** shows the evolution of the objective function (the current best value) in the optimization process. Although the objective function still shows some improvements after approximately 110 iterations, the incremental improvement observed in F is no longer substantial. Because of the high computational cost involved in the OUU process, additional iterations were not possible. Therefore, the optimization result remains as an approximate optimal solution. In addition, although it is favorable to conduct the GP optimization routine multiple times to decrease the randomness effect in the process, the computational cost of the OUU would not simply allow us to easily conduct such tasks in costly simulation problems.

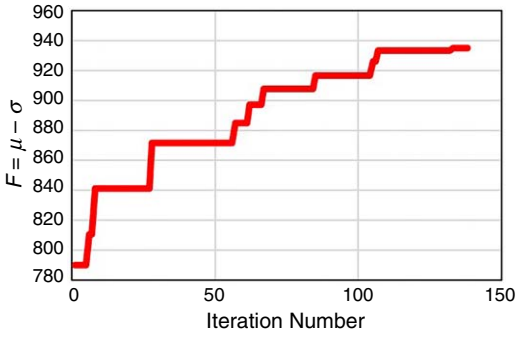

Fig. 12—Current best value of *F* at each iteration of the optimization under uncertainty.

Table 4 provides the optimal values and the ranges of CVs used in the optimization runs. The ranges of variables have been determined by consulting the field engineers responsible for implementing the HNP. For example, the maximum injection pressure is below the fracturing pressure to avoid creating unwanted fractures during the gas-injection process. The maximum injection rate is also dictated by the current maximum compressor power available. The duration of production and soaking periods are selected wide enough to ensure different possibilities. Compared with our previous study (Hamdi et al. 2018a), the maximum possible injection rate is significantly decreased to 18.9 Mscf/D (equivalent to approximately 5 MMscf/D for the whole well). This is to assess the feasibility of the HNP gas injection with lower gas-injection rates. The results demonstrate that the lower confidence interval is maximized by decreasing the soaking time to its minimum value and increasing the injection rate to its maximum value. In addition, the optimal injection and production periods are estimated as 40 and 70 days, respectively, while the producing bottomhole pressure is above the bubblepoint pressure. **Fig. 13** shows the corresponding probabilistic forecast for the suggested CVs. The figure indicates that the median (P50) of the probabilistic forecast from cyclic rich gas injection is 30% higher than the median of recovery from natural depletion with a constant bottomhole pressure of 1,500 psia (i.e., the last recorded pressure of historical data). This is in agreement with the EOG-reported 30% to 70% increase in recovery by implementing their cyclic gas injection method (Rassenfoss 2017). In addition, a lag time of approximately 1 year is apparent before the median of the HNP gas-injection recovery curve starts to pick up (Fig. 13). The probabilistic oil rate prediction is shown in **Fig. 14,** demonstrating a noticeable increase in the oil production rate when the HNP rich gas-injection process is conducted.

	Minimum	Maximum	Optimum
Injection rate (q_{inj}), Mscf/D	1.89	18.9	18.9
Injection time (t_{inj}), day	14	180	40
Soaking time (t_{soak}), day	1	90	1
Production pressure (p_{prod}), psi	1,500	6,000	4,140
Production time (t_{prod}), day	14	360	70
Maximum injection pressure	Initial pressure (8,700 psi)		
Gas production rate	Maximum gas-injection rate (18.9 Mscf/D)		

Table 4—Range of control variables (operational conditions) used in the optimization process and the optimum parameters that lead to the highest oil recovery. These values correspond to the one-quarter fracture model. To scale up to the full well results, volumes must be multiplied by 264.

Our optimization results are based on an objective function, which does not include any economic considerations to impose additional constraints on the HNP operation such as the number of cycles. Having an additional constraint on the number of cycles could potentially impact the influence of some parameters such as the soaking time. For instance, although increasing the soaking time might increase the recovery in the first few cycles (Gamadi et al. 2014), it cannot offset the recovery loss caused by shutting the well if we are able to operate the process with no limit on the number of cycles (Gamadi et al. 2014; Sheng 2017).

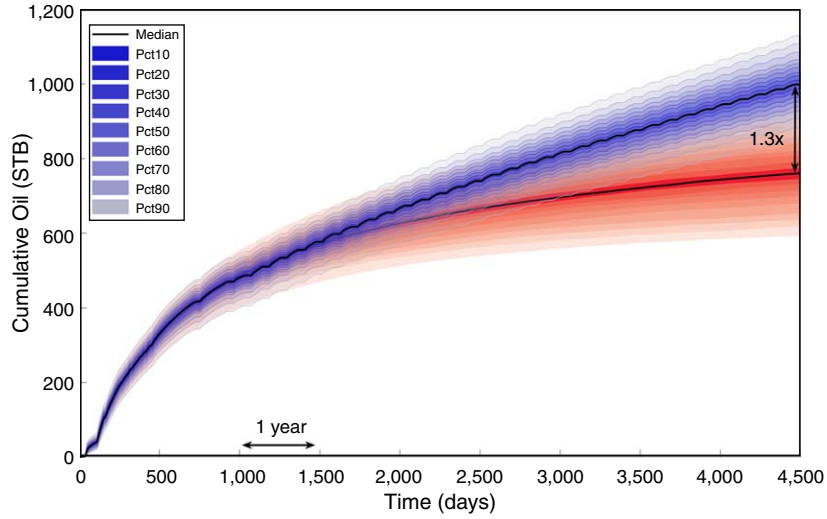

Fig. 13—Fan chart representing the distribution around the median for oil recovery predicted using primary and enhanced recovery processes. The blue curves are the probabilistic prediction of cumulative oil production (percentiles around median) obtained from cyclic rich gas injection using the optimized CVs. The red curves represent the probabilistic prediction of primary depletion using a constant bottomhole flowing pressure of 1,500 psi.

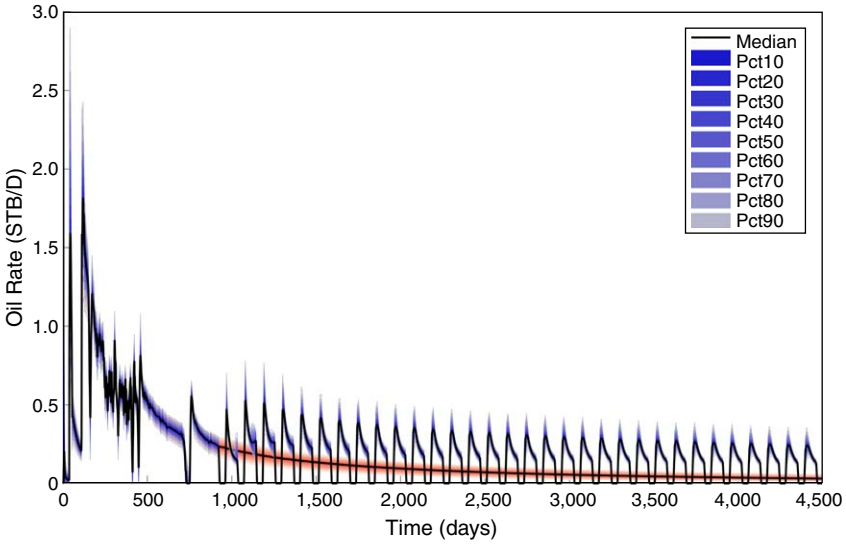

Fig. 14—Fan chart representing the distribution around median for oil production rates using primary and enhanced recovery processes. The blue curves are the probabilistic prediction of oil rate obtained from cyclic rich gas injection using the optimized CVs. The red curves represent the probabilistic prediction of primary depletion using a constant bottomhole pressure of 1,500 psi.

Discussion

Other Optimization Solutions. The purpose of the optimization routine is to find a set of most likely parameters that can maximize the objective function. A drawback of this approach is that the solutions that are only slightly worse than the optimum case will be ignored. This situation is exacerbated when the objective function is nearly flat in some areas of the domain. Therefore, it is worth investigating the shape of the objective function to identify the areas near the global optima with favorable solutions. A benefit of using the GP-based optimization method is that it can provide an accurate proxy to mimic the behavior of the objective function without running the numerical simulation. **Fig. 15** shows the negative lower confidence interval ($-F$) as a function of two variables, whereas the rest of the parameters are set to their optimal values. The figure shows that with other production scenarios such as $3{,}500 < p_{\text{prod}} < 4{,}500$, $50 < t_{\text{prod}} < 120$ days, $20 < t_{\text{inj}} < 70$ days, $0 < t_{soak} < 7$ days, and $16 < q_{\text{inj}} < 18.9$ Mscf/D, one can still achieve competitive solutions with high oil recovery. These ranges provide the operator with operational flexibility to implement the HNP in practice.

Lean/Rich Gas Injection. **Fig. 16** compares the cumulative oil production and oil rate from the primary depletion simulation with constant bottomhole pressure ($p_{wf} = 1{,}500$ psi) with that obtained from cyclic lean and rich gas-injection scenarios (the gas compositions are listed in Table 3). The numerical simulation model used for this comparison corresponds to the best solution case with the highest oil recovery obtained from the results of OUU. The results demonstrate that the cyclic rich and lean gas injection scenarios could increase oil recovery by 70% and 20% over primary production, respectively. This illustrates the superiority of rich gas for extracting oil in this volatile oil Duvernay shale example.

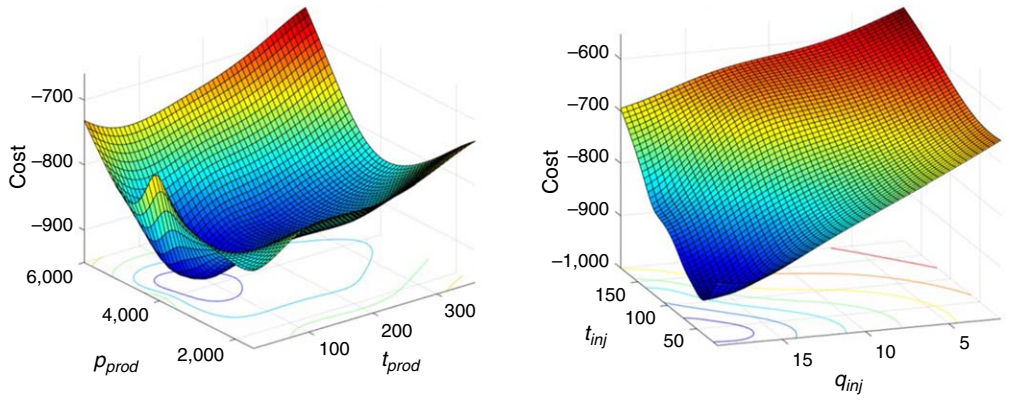

Fig. 15—Negative lower confidence interval (cost) after 10 years of conducting cyclic rich gas injection as a function of production pressure and production time (left) and injection rate and injection time (right). The other parameters are fixed at their optimal values. The cost function is multiplied by −1 to represent a minimization problem in the OUU process.

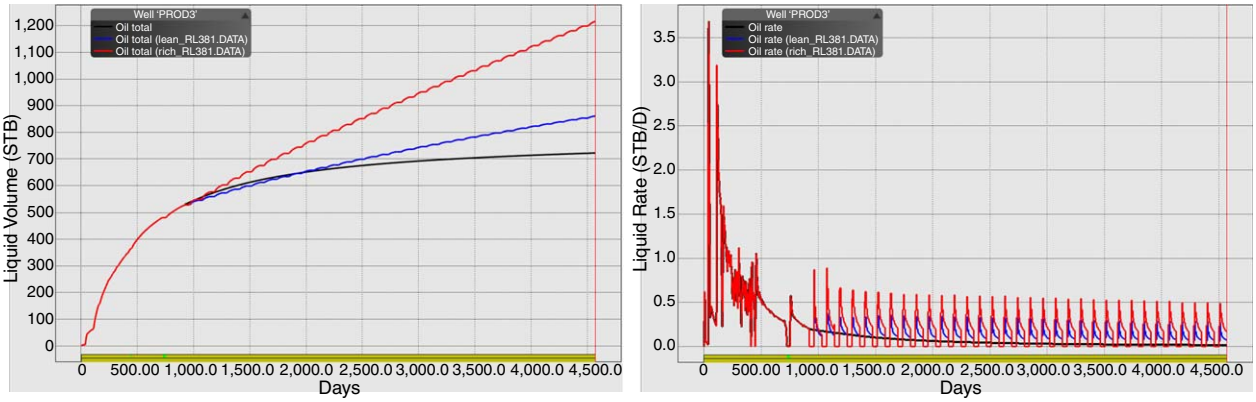

Fig. 16—Comparison between cumulative oil production (left) and oil production rate (right) from the primary depletion (with a constant pressure of 1,500 psi) and that obtained from cyclic lean and rich gas-injection processes. The black curve is primary depletion, whereas the blue and red curves represent the cyclic lean and rich gas-injection scenarios, respectively.

The efficiency of rich gas in extracting the oil is attributed to the way that the injectant interacts with the reservoir oil. One of the important recovery mechanisms in the gas injection process is mass transfer. If the injected gas is rich in intermediate hydrocarbon components (e.g., C_2 to C_6), the recovery mechanism is referred to as condensing gas drive, whereas for the lean gas injection, the process it is called vaporizing gas drive. Although, in reality, a combination of these two mechanisms work together, it is still important to understand how the two contacting fluids exchange components to approach equilibrium. This can help to better understand the results of the gas injection process. With a purely vaporizing gas drive mechanism, the injectant contacts the in-situ oil and extracts the intermediate hydrocarbon components from the oil. Because the gas has a higher mobility than the in-situ fluid, it will propagate further into the formation as injection continues. The moving gas front gets richer with time until eventually miscibility happens at the leading edge, where the composition of the oil and the gas becomes very similar. However, as the new injected gas continuously passes the in-situ oil at the tail, the remaining oil, which is close to the injection point, becomes heavier with injection. The outcome of this mass transfer mechanism is that during the HNP gas injection, the producing oil, which is largely sourced from the near fracture area, tends to become heavier with time.

Conversely, with a purely condensing gas drive mechanism, when the injectant meets the reservoir oil, some of the intermediate hydrocarbon components in the gas condense into the in-situ oil. Therefore, the gas that has lost part of its intermediate components becomes leaner while it propagates into the formation. The leading front of the gas eventually becomes very lean as it contacts fresh oil in the reservoir. However, at the tail, closer to the injection point, the compositions of the injectant and the in-situ oil approach equilibrium (i.e., miscibility). The implication of this mass transfer process is that, during HNP gas injection, the producing oil becomes lighter as the injection continues. A thorough discussion of mass transfer mechanisms happening during miscible gas injection is given in Khabibullin et al. (2017). **Fig. 17** compares the density of the produced fluid from primary depletion with that from cyclic gas-injection processes. As expected, for the rich gas-injection scenario, the produced oil is lighter than for lean gas injection. This can be also observed from the variation of oil mole fractions with production **(Fig. 18).** Fig. 18 illustrates that HNP with rich gas injection produces greater mole fractions of intermediate components (C_2 to C_7; Fig. 18, upper plots) and less mole fractions of heavy components (C_8 to C_{20+}; Fig. 18, lower plots) in each cycle. Conversely, lean gas injection leads to the production of heavier components.

Overall Sensitivity Analysis. Having a large pool of samples obtained from the optimization runs, which includes the variability in both UVs and CVs, higher-order sensitivity analysis can now be performed. Using some special surrogate models such as polynomial chaos expansion (PCE) proxies enables one to correctly account for the distribution of the parameters in the sensitivity analysis and the propagation of uncertainty. **Fig. 19** provides an approximation of total sensitivity indices that are obtained by Sobol' decomposition of the fitted PCE model (Sudret 2008; Marelli and Sudret 2014). This figure demonstrates that, in general, the design variables (CVs) have a larger impact on HNP performance than the UVs. However, some UVs such as the permeability multiplier, $x_{f\text{-eff}}$, molecular diffusion coefficients, and fracture permeability can also considerably influence the HNP performance. Understanding the impact of UVs on the

HNP performance is very important for a feasibility study because, although some UVs such as diffusion coefficients may not show a large sensitivity to primary production data and history-matching results, their impact is magnified for the subsequent cyclic gas-injection processes.

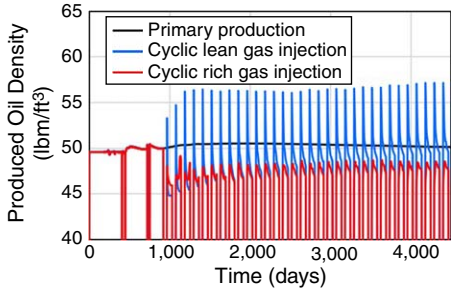

Fig. 17—Comparison between the produced oil density from primary depletion (with a constant pressure of 1,500 psi) and that from cyclic lean and rich gas-injection processes.

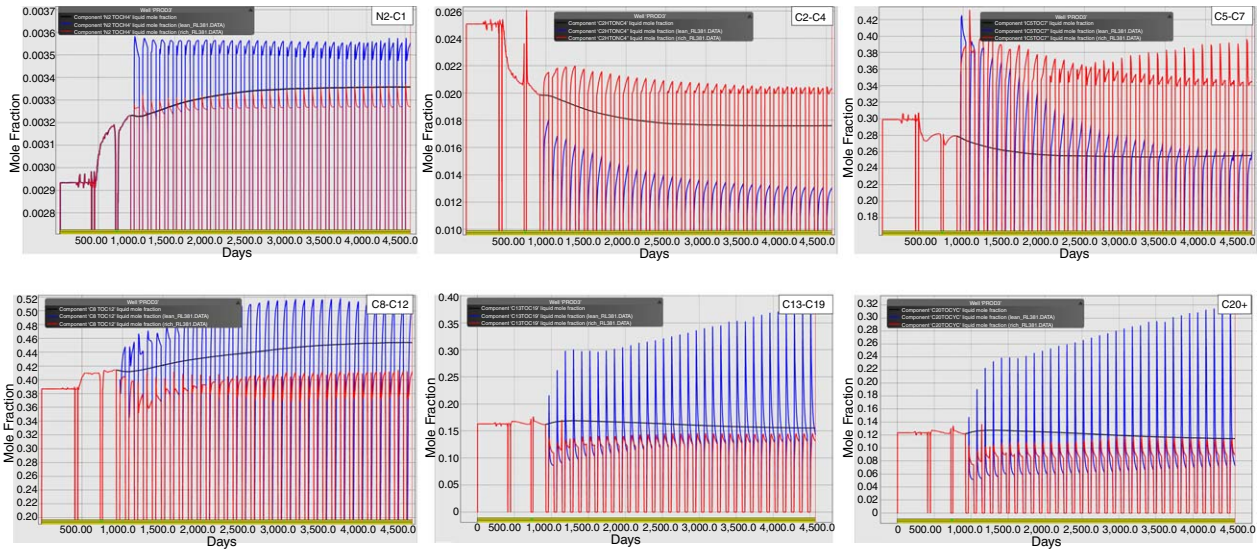

Fig. 18—Comparison between the mole fraction of produced oil from primary depletion (with a constant pressure of 1,500 psi) and that from cyclic lean and rich gas-injection processes. The black curve is for primary depletion, whereas the blue and red curves represent cyclic lean and rich gas injection, respectively.

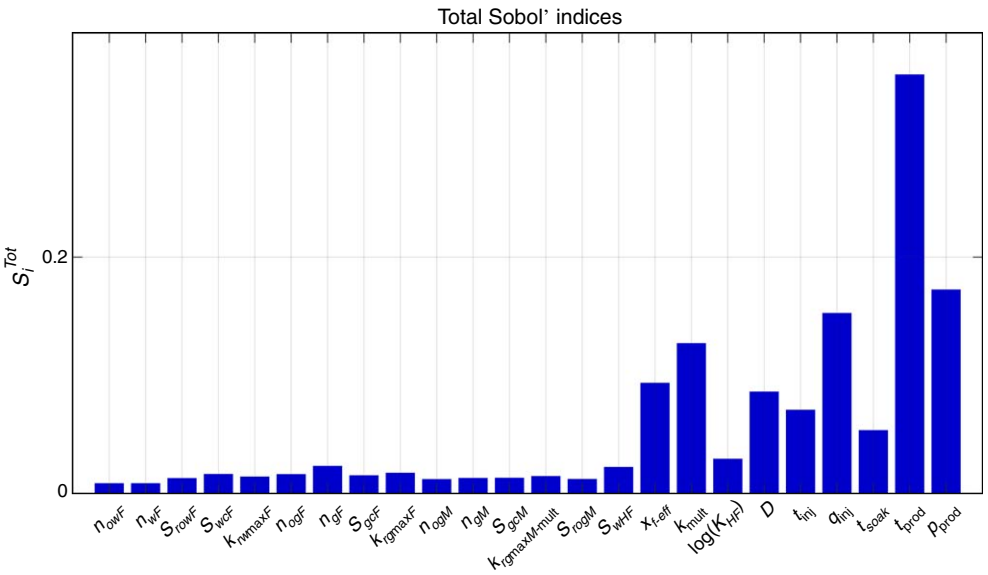

Fig. 19—Sensitivity of HNP performance with respect to all uncertain and control variables. The Sobol' sensitivity indices are obtained by repetitive sampling from the surrogate model (i.e., polynomial chaos expansion).

Conclusions

In this work, the impact of uncertainty on the design of optimum operational conditions for the cyclic gas-injection process is investigated. An MFHW completed in the volatile oil window of the Duvernay shale in Canada is used as an example. An LEFR numerical model was selected to conduct an initial sensitivity analysis to identify the most sensitive parameters. An adaptive sampling method (LOLA-Voronoi) and a multidimensional surrogate modeling approach (kriging) were then used to conduct Bayesian history matching to obtain the posterior distributions. The posterior distributions and other un- (less) impactful parameters were sampled to conduct the OUU. The lower confidence interval was optimized with the maximum gas-injection rate, the minimum soaking time, and a bottomhole pressure close to the bubblepoint pressure of the in-situ fluid. The injection and production time of the HNP process were optimized at 40 and 70 days, respectively. The results demonstrated that, for the studied case, cyclic rich gas injection could lead to an increase in oil recovery over primary production, with a P50 (median) that is 1.3 times larger than the median of an extended primary depletion period with a constant bottomhole pressure of 1,500 psi. Using the GP-based optimization method, it was possible to obtain an accurate surrogate model to explore various areas of the objective function to identify other competitive solutions that could still result in high HNP incremental oil recoveries. This is a very important point that can lead the operator to explore more effective and efficient surface facility and EOR scheme configurations. It was also observed that lean gas injection could increase oil recovery, but to a lesser extent. This was attributed to a different mass transfer mechanism than happens during lean gas injection. In addition, it was shown that, although CVs generally have a larger impact on HNP performance, some UVs such as diffusion coefficients could also considerably affect the cumulative oil production from the HNP process.

Nomenclature

$\boldsymbol{C}_D$ = covariance matrix
C_o, C_e = odd and even order effects
D = data or diffusion coefficients
D_o, D_g = diffusion coefficients for the oil and gas components
E = combined measurement error and simulation inaccuracy
F = lower confidence interval objective function
$_{\text{inj}}$ = injection
k = matrix (formation) permeability, md
k_{rg}, k_{ro}, k_{rw} = gas, oil, and water relative permeability, respectively
$k_{rg\max}, k_{rw\max}$ = maximum gas and water relative permeability values
K_{mult} = permeability multiplier
L = likelihood
$\boldsymbol{m}$ = model parameters
m = measured data
M, F = matrix and fracture; subscripts
n_{og}, n_g = Corey's exponent for oil and gas relative permeability curves
n_{ow}, n_w = Corey's exponent for oil and water relative permeability curves
p = pressure, psia
p_i = initial pressure, psia
$_{\text{prod}}$ = production
p_{wf} = flowing bottomhole pressure, psia
Pr = probability distribution function
q_g, q_o, q_w = surface flow rate for gas, Mscf/D; oil, STB/D; and water, STB/D, respectively
$S(i)$ = sensitivity index for parameter i
S_g, S_o, S_w = gas, oil, and water saturations, respectively
S_{gc}, S_{wc} = critical gas and water saturations, respectively
S_{rog}, S_{row} = residual oil saturation in the presence of gas or water
$_{\text{sim}}$ = simulated data
t = time, days
x_f = fracture half-length, ft
$x_{f\text{-eff}}$ = effective fracture half-length (LEFR extend), ft
μ = mean or average value
σ = standard deviation
$\sigma_o, \sigma_w, \sigma_g, \sigma_p$ = measurement errors in oil, water, and gas rates and pressure

Acknowledgments

Hamidreza Hamdi thanks Rock Flow Dynamics for supporting his independent research. Chris Clarkson acknowledges Ovintiv and Shell for support of his chair position in Unconventional Gas and Light Oil research at the University of Calgary, Department of Geoscience. The authors further thank the sponsors of the Tight Oil Consortium, hosted at the University of Calgary, for support. Partial support for this work was provided through an NSERC CRD grant held by Clarkson. The authors also thank Ivo Couckuyt of Ghent University-iMinds and Charles Tong of Lawrence Livermore National Laboratory for great discussion and help on statistical modeling and analysis.

References

Aanonsen, S. I., Aavatsmark, I., Barkve, T. et al. 2003. Effect of Scale Dependent Data Correlations in an Integrated History Matching Loop Combining Production Data and 4D Seismic Data. Paper presented at the SPE Reservoir Simulation Symposium, Houston, Texas, USA, 3–5 February. SPE-79665-MS. https://doi.org/10.2118/79665-MS.

Alfarge, D., Wei, M., and Bai, B. 2017a. Factors Affecting CO_2-EOR in Shale-Oil Reservoirs: Numerical Simulation Study and Pilot Tests. *Energy & Fuels* **31** (8): 8462–8480. https://doi.org/10.1021/acs.energyfuels.7b01623.

Alfarge, D., Wei, M., and Bai, B. 2017b. IOR Methods in Unconventional Reservoirs of North America: Comprehensive Review. Paper presented at the SPE Western Regional Meeting, Bakersfield, California, USA, 23–27 April. SPE-185640-MS. https://doi.org/10.2118/185640-MS.

Ampomah, W., Balch, R. S., Cathar, M. et al. 2016. Performance of CO_2-EOR and Storage Processes under Uncertainty. Paper presented at the SPE Europec/EAGE Conference and Exhibition, Vienna, Austria, 30 May–2 June. SPE-180084-MS. https://doi.org/10.2118/180084-MS.

Badru, O. and Kabir, C. S. 2003. Well Placement Optimization in Field Development. Paper presented at the SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA, 5–8 October. SPE-84191-MS. https://doi.org/10.2118/84191-MS.

Bailey, W. J. and Benoıt, C. 2011. Forecast Optimization and Value of Information under Uncertainty. In *Uncertainty Analysis and Reservoir Modeling: AAPG Memoir 96*, ed. Y. Zee Ma and Paul R. La Pointe. Tulsa, Oklahoma, USA: AAPG.

Bard, Y. 1974. *Nonlinear Parameter Estimation*. New York, New York, USA: Academic Press.

Blatman, G. 2009. *Adaptive Sparse Polynomial Chaos Expansions for Uncertainty Propagation and Sensitivity Analysis*. PhD dissertation, Blaise Pascal University, Clermont-Ferrand, France.

Buhmann, M. D. 2003. *Radial Basis Functions: Theory and Implementations*. Cambridge, UK: Cambridge University Press.

Chen, Y., Oliver, D. S., and Zhang, D. 2009. Efficient Ensemble-Based Closed-Loop Production Optimization. *SPE J*. **14** (4): 634–645. SPE-112873-PA. https://doi.org/10.2118/112873-PA.

Chiles, J. P. and Delfiner, P. 2012. *Geostatistics: Modeling Spatial Uncertainty*, Vol. 713. Hoboken, New Jersey, USA: John Wiley & Sons.

Christiansen, R. L. 2007. Relative Permeability and Capillary Pressure. In *Petroleum Engineering Handbook: Vol. 1: General Engineering*, ed. John R. Fanchi, Chap. 15, 727–765. Richardson, Texas, USA: Society of Petroleum Engineers.

Christie, M. A. and Bazargan, H. 2012. Efficient Polynomial Chaos Proxy-Based History Matching and Uncertainty Quantification for Complex Geological Structures. Paper presented at the SPE Kuwait International Petroleum Conference and Exhibition, Kuwait City, Kuwait, 10–12 December. SPE-163282-MS. https://doi.org/10.2118/163282-MS.

Chugunov, N., Ramakrishnan, T. S., Lukyanov, A. et al. 2015. Method for Adaptive Optimization of EOR Performance under Uncertainty. Paper presented at the SPE Reservoir Simulation Symposium, Houston, Texas, USA, 23–25 February. SPE-173295-MS. https://doi.org/10.2118/173295-MS.

Cotter, S. C. 1979. A Screening Design for Factorial Experiments with Interactions. *Biometrika* **66** (2): 317–320. https://doi.org/10.1093/biomet/66.2.317.

Crombecq, K. 2011. *Surrogate Modelling of Computer Experiments with Sequential Experimental Design*. PhD dissertation, Universiteit Gent, Gent, Belgium.

Crombecq, K., Tommasi, L. D., Gorissen, D. et al. 2009. A Novel Sequential Design Strategy for Global Surrogate Modeling. Paper presented at the 2009 Winter Simulation Conference, Austin, Texas, USA, 13–16 December 2009. https://doi.org/10.1109/WSC.2009.5429687.

Downey, A. B. 2013. *Think Bayes: Bayesian Statistics in Python*. Sebastopol, California, USA: O'Reilly Media, Inc.

Fai-Yengo, V., Rahnema, H., and Alfi, M. 2014. Impact of Light Component Stripping during CO_2 Injection in Bakken Formation. Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, USA, 25–27 August. URTEC-1922932-MS. https://doi.org/10.15530/URTEC-2014-1922932.

Fothergill, P., Boskovic, D., Schoellkopf, N. et al. 2014. Regional Modelling of the Late Devonian Duvernay Formation, Western Alberta, Canada. Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, USA, 25–27 August. URTEC-1923935-MS. https://doi.org/10.15530/URTEC-2014-1923935.

Gamadi, T. D., Sheng, J. J., and Soliman, M. Y. 2013. An Experimental Study of Cyclic Gas Injection To Improve Shale Oil Recovery. Paper presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA, 30 September–2 October. SPE-166334-MS. https://doi.org/10.2118/166334-MS.

Gamadi, T. D., Sheng, J. J., Soliman, M. Y. et al. 2014. An Experimental Study of Cyclic CO_2 Injection To Improve Shale Oil Recovery. Paper presented at the SPE Improved Oil Recovery Symposium, Tulsa, Oklahoma, USA, 12–16 April. SPE-169142-MS. https://doi.org/10.2118/169142-MS.

Gamerman, D. 1997. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Boca Raton, Florida, USA: Chapman & Hall/CRC Press LLC.

Geman, S. and Geman, D. 1984. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Trans PAMI* **6** (6): 721–741. https://doi.org/10.1109/TPAMI.1984.4767596.

Gong, W., Cai, Z., and Wang, Y. 2014. Repairing the Crossover Rate in Adaptive Differential Evolution. *Applied Soft Computing* **15:** 149–168. https://doi.org/ 10.1016/j.asoc.2013.11.005.

Goodwin, N. 2015. Bridging the Gap between Deterministic and Probabilistic Uncertainty Quantification Using Advanced Proxy Based Methods. Paper presented at the SPE Reservoir Simulation Symposium, Houston, Texas, USA, 23–25 February. SPE-173301-MS. https://doi.org/10.2118/173301-MS.

Gorissen, D., Couckuyt, I., Demeester, P. et al. 2010. A Surrogate Modeling and Adaptive Sampling Toolbox for Computer Based Design. *J Mach Learn Res* **11:** 2051–2055.

Hamdi, H., Behmanesh, H., Clarkson, C. R. et al. 2015. Using Differential Evolution for Compositional History-Matching of a Tight Gas Condensate Well in the Montney Formation in Western Canada. *J Natural Gas Sci & Eng* **26:** 1317–1331. https://doi.org/10.1016/j.jngse.2015.08.015.

Hamdi, H., Clarkson, C. R., Ghanizadeh, A. et al. 2018a. Huff-N-Puff Gas Injection Performance in Shale Reservoirs: A Case Study from Duvernay Shale in Alberta, Canada. Presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas, USA, 23-25 July URTEC-2902835-MS. https://doi.org/10.15530/URTEC-2018-2902835.

Hamdi, H., Couckuyt, I., Costa Sousa, M. et al. 2017a. Gaussian Processes for History-Matching: Application to an Unconventional Gas Reservoir. *Comput Geosci* **21** (2): 267–287. https://doi.org/10.1007/s10596-016-9611-2.

Hamdi, H., Couckuyt, I., Dhaene, T. et al. 2018b. Efficient Multi-Objective History-Matching Using Gaussian Processes. Paper presented at the ECMOR XVI—16th European Conference on the Mathematics of Oil Recovery, Barcelona, Spain, 3 September 2018. https://doi.org/10.3997/2214-4609.201802146.

Hamdi, H., Sousa, M. C., and Behmanesh, H. 2017b. Bayesian History-Matching and Probabilistic Forecasting for Tight and Shale Wells. Paper presented at the SPE Unconventional Resources Conference, Calgary, Alberta, Canada, 15–16 February SPE-185082-MS. https://doi.org/10.2118/185082-MS.

Hao, S., Dengen, Z., Adwait, C. et al. 2016. Quantifying Shale Oil Production Mechanisms by Integrating a Delaware Basin Well Data from Fracturing to Production. Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, San Antonio, Texas, USA, 1–3 August. URTEC-2425721-MS. https://doi.org/10.15530/URTEC-2016-2425721.

Hoffman, B. T. 2018. Huff-N-Puff Gas Injection Pilot Projects in the Eagle Ford. Paper presented at the SPE Canada Unconventional Resources Conference, Calgary, Alberta, Canada, 13–14 March. SPE-189816-MS. https://doi.org/10.2118/189816-MS.

Hoffman, T. and Evans, J. G. 2016. Improved Oil Recovery IOR Pilot Projects in the Bakken Formation. Paper presented at the SPE Low Perm Symposium, Denver, Colorado, USA, 5–6 May. SPE-180270-MS. https://doi.org/10.2118/180270-MS.

Jones, D. 2001. A Taxonomy of Global Optimization Methods Based on Response Surfaces. *J Global Optim* **21** (4): 345–383. https://doi.org/10.1023/A:1012771025575.

Jones, D. R., Schonlau, M., and Welch, W. J. 1998. Efficient Global Optimization of Expensive Black-Box Functions. *J Global Optim* **13** (4): 455–492. https://doi.org/10.1023/a:1008306431147.

Khabibullin, R., Emadi, A., Abu Grin, Z. et al. 2017. Investigation of CO_2 Application for Enhanced Oil Recovery in a North African Field—A New Approach to EOS Development. Paper presented at the IOR 2017—19th European Symposium on Improved Oil Recovery, Stavanger, Norway, 24–27 April. https://doi.org/10.3997/2214-4609.201700276.

Kruschke, J. 2014. *Doing Bayesian Data Analysis*, 2nd ed. New York, New York, USA: Academic Press.

Lataniotis, C., Marelli, S., and Sudret, B. 2015. UQLab User Manual—Kriging (Gaussian Process Modelling). Report UQLab-V0.9-105, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich, Switzerland.

Li, L. and Sheng, J. J. 2016. Experimental Study of Core Size Effect on CH_4 Huff-n-Puff Enhanced Oil Recovery in Liquid-Rich Shale Reservoirs. *J Natural Gas Sci & Eng* **34:** 1392–1402. https://doi.org/10.1016/j.jngse.2016.08.028.

Lyster, S., Corlett, H. J., and Berhane, H. 2017. Hydrocarbon Resource Potential of the Duvernay Formation in Alberta–Update. AER/AGS Open File Report 2017-02, Alberta Energy Regulator, Edmonton, Alberta, Canada (July 2017).

Marelli, S. and Sudret, B. 2014. UQLab: A Framework for Uncertainty Quantification in Matlab. Paper presented at the 2nd International Conference on Vulnerability, Risk Analysis and Management and the 6th International Symposium on Uncertainty, Modeling, and Analysis, Liverpool, UK, 13–16 July. https://doi.org/10.1061/9780784413609.257.

Mariethoz, P. G. and Caers, P. J. 2014. *Multiple-Point Geostatistics: Stochastic Modeling with Training Images*. New York, New York, USA: Wiley.

Marmin, S., Chevalier, C., and Ginsbourger, D. 2015. Differentiating the Multipoint Expected Improvement for Optimal Batch Design. In *Machine Learning, Optimization, and Big Data: First International Workshop, MOD 2015*, ed. Panos Pardalos, Mario Pavone, Giovanni Maria Farinella, and Vincenzo Cutello, 37–48. Cham, Switzerland: Springer International Publishing.

Mehranfar, R., Marquez, L., Altman, R. et al. 2018. Optimization under Uncertainty for Reliable Unconventional Play Evaluation. A Case Study in Vaca Muerta Shale Gas Blocks, Argentina. Paper presented at the SPE Trinidad and Tobago Section Energy Resources Conference, Port of Spain, Trinidad and Tobago, 25–26 June. SPE-191272-MS. https://doi.org/10.2118/191272-MS.

Mohammad, Y. and Nishida, T. 2016. *Data Mining for Social Robotics: Toward Autonomously Social Robots*. New York, New York, USA: Springer International Publishing.

Morales, A. N., Nasrabadi, H., and Zhu, D. 2011. A New Modified Genetic Algorithm for Well Placement Optimization under Geological Uncertainties. Paper presented at the SPE Europec/EAGE Annual Conference and Exhibition, Vienna, Austria, 23–26 May. SPE-143617-MS. https://doi.org/10.2118/143617-MS.

Murphy, K. P. 2012. *Machine Learning: A Probabilistic Perspective*. Cambridge, Massachusetts, USA: The MIT Press.

Nguyen, N. T. B., Dang, C. T. Q., Nghiem, L. X. et al. 2016. Robust Optimization of Unconventional Reservoirs under Uncertainties. Paper presented at the SPE Europec/EAGE Conference and Exhibition, Vienna, Austria, 30 May–2 June. SPE-180108-MS. https://doi.org/10.2118/180108-MS.

Picheny, V., Wagner, T., and Ginsbourger, D. 2013. A Benchmark of Kriging-Based Infill Criteria for Noisy Optimization. *Struct & Multidisc Optim* **48** (3): 607–626. https://doi.org/10.1007/s00158-013-0919-4.

Podhoretz, S. B. and Valkó, P. P. 2014. Hydraulic Fracture Design for the Lower Tertiary Gulf of Mexico: Optimization under Uncertainty. Paper presented at the Offshore Technology Conference, Houston, Texas, USA, 5–8 May. OTC-25179-MS. https://doi.org/10.4043/25179-MS.

Privault, N. 2013. *Understanding Markov Chains: Examples and Applications*. Singapore: Springer.

Pu, W. and Hoffman, B. T. 2014. EOS Modeling and Reservoir Simulation Study of Bakken Gas Injection Improved Oil Recovery in the Elm Coulee Field, Montana. Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, USA, 25–27 August. URTEC-1922538-MS. https://doi.org/10.15530/URTEC-2014-1922538.

Rasmussen, C. E. and Williams, C. K. I. 2005. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. Cambridge, Massachusetts, USA: The MIT Press.

Rassenfoss, S. 2017. Shale EOR Works, but Will It Make a Difference. *J Pet Technol* **69** (10): 34–40. https://doi.org/10.2118/1017-0034-JPT.

Robinson, D. B. and Peng, D.-Y. 1978. *The Characterization of the Heptanes and Heavier Fractions for the GPA Peng-Robinson Programs*. Tulsa, Oklahoma, USA: Gas Processors Association.

Rock Flow Dynamics. 2018. *tNavigator Reservoir Simulator's User Manual (v.18.3),* Houston, Texas, USA.

Saltelli, A., Ratto, M., Andres, T. et al. 2008. *Global Sensitivity Analysis: The Primer*. 304 p. John Wiley & Sons, Ltd. https://onlinelibrary.wiley.com/doi/book/10.1002/9780470725184.

Schlumberger. 2011. Eclipse V.2011 software, Houston, Texas, USA.

Schulze-Riegert, R., Nwakile, M., Skripkin, S. et al. 2016. Scalability and Performance Efficiency of History Matching Workflows Using MCMC and Adjoint Techniques Applied to the Norne North Sea Reservoir Case Study. Paper presented at the SPE Europec/EAGE Conference and Exhibition, Vienna, Austria, 30 May–2 June. SPE-180106-MS. https://doi.org/10.2118/180106-MS.

Sheng, J. J. 2017. Optimization of Huff-n-Puff Gas Injection in Shale Oil Reservoirs. *Petroleum* **3** (4): 431–437. https://doi.org/10.1016/j.petlm.2017.03.004.

Slotte, P. A. and Smorgrav, E. 2008. Response Surface Methodology Approach for History Matching and Uncertainty Assessment of Reservoir Simulation Models. Paper presented at the SPE Europec/EAGE Conference and Exhibition, Rome, Italy, 9–12 June. SPE-113390-MS. https://doi.org/10.2118/113390-MS.

Snoek, J., Larochelle, H., and Adams, R. 2012. Practical Bayesian Optimization of Machine Learning Algorithms. http://arxiv.org/abs/1206.2944.

Sobol', I. M. 1967. On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals. *USSR Comput Math & Math Physics* **7** (4): 86–112. https://doi.org/10.1016/0041-5553(67)90144-9.

Song, C. and Yang, D. 2013. Performance Evaluation of CO_2 Huff-n-Puff Processes in Tight Oil Formations. Paper presented at the SPE Unconventional Resources Conference Canada, Calgary, Alberta, Canada, 5–7 November. SPE-167217-MS. https://doi.org/10.2118/167217-MS.

Srivastava, M. K., Khan, A. H., and Srivastava, N. 2014. *Statistical Inference: Theory of Estimation*. Delhi, India: PHI Learning Private Limited.

Srinivas, N., Krause, A., Kakade, S. M. et al. 2010. Gaussian Process Optimization in the Bandit Setting: No Regret and Experimental Design. https://arxiv.org/abs/0912.3995.

Stein, M. 1999. *Interpolation of Spatial Data: Some Theory for Kriging*. New York, New York, USA: Springer Science & Business Media.

Storn, R. and Price, K. 1995. *Differential Evolution—A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces*. Berkeley, California, USA: International Computer Science Institute.

Sudret, B. 2008. Global Sensitivity Analysis Using Polynomial Chaos Expansions. *Reliability Eng & System Safety* **93** (7): 964–979. https://doi.org/10.1016/j.ress.2007.04.002.

Tilke, P., Zhou, W., Wang, Y. et al. 2015. Automated Field Development Planning for Unconventional Shale Gas and Tight Oil. Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, San Antonio, Texas, USA, 20–22 July. URTEC-2169844-MS. https://doi.org/10.15530/URTEC-2015-2169844.

Tong, C. 2018. *Psuade*. Livermore, California, USA: Center for Applied Scientific Computing Lawrence Livermore National Laboratory.

Wan, T. 2013. *Evaluation of the EOR Potential in Shale Oil Reservoirs by Cyclic Gas Injection*. Master's thesis, Texas Tech University, Lubbock, Texas, USA.

Wantawin, M., Yu, W., Dachanuwattana, S. et al. 2017a. An Iterative Response-Surface Methodology by Use of High-Degree-Polynomial Proxy Models for Integrated History Matching and Probabilistic Forecasting Applied to Shale-Gas Reservoirs. *SPE J.* **22** (6): 2012–2031. SPE-187938-PA. https://doi.org/10.2118/187938-PA.

Wantawin, M., Yu, W., and Sepehrnoori, K. 2017b. 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 Res Eval & Eng* **20** (3): 613–626. SPE-185181-PA. https://doi.org/10.2118/185181-PA.

Wilson, A. G. and Adams, R. P. 2013. Gaussian Process Kernels for Pattern Discovery and Extrapolation. https://arxiv.org/abs/1302.4245.

Wolpert, D. H. and Macready, W. G. 1997. No Free Lunch Theorems for Optimization. *Trans Evol Comp* **1** (1): 67–82. https://doi.org/10.1109/4235.585893.

Yang, C., Nghiem, L., Erdle, J. et al. 2015. An Efficient and Practical Workflow for Probabilistic Forecasting of Brown Fields Constrained by Historical Data. Paper presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, USA, 28–30 September. SPE-175122-MS. https://doi.org/10.2118/175122-MS.

Ye, Q. 1998. *Latin Hypercube Designs for Computer Experiments*. PhD dissertation, University of Michigan, Ann Arbor, Michigan, USA.

Yeh, T.-H., Jimenez, E., Van Essen, G. et al. 2014. Reservoir Uncertainty Quantification Using Probabilistic History Matching Workflow. Paper presented at the SPE Annual Technical Conference and Exhibition, Amsterdam, The Netherlands, 27–29 October. SPE-170893-MS. https://doi.org/10.2118/170893-MS.

Yu, W., Lashgari, H., and Sepehrnoori, K. 2014. Simulation Study of CO_2 Huff-n-Puff Process in Bakken Tight Oil Reservoirs. Paper presented at the SPE Western North American and Rocky Mountain Joint Meeting, Denver, Colorado, USA, 17–18 April. SPE-169575-MS. https://doi.org/10.2118/169575-MS.