QUANTIFYING THE BENEFIT OF ENSEMBLE SELECTION

FOR OPTIMIZATION UNDER UNCERTAINTY

A Thesis

by

CHANAPOL BHURIPANYO

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Duane A. McVay
Co-Chair of Committee,	Eduardo Gildin
Committee Members,	Thomas A. Blasingame
Head of Department,	A. Daniel Hill

August 2014

Major Subject: Petroleum Engineering

ABSTRACT

Due to significant uncertainty in reservoir parameters, maximizing reservoir potential is an extremely difficult task. To be able to make decisions that maximize the reservoir potential, knowledge of possible ranges of reservoir parameters and production optimization are critical. The closed-loop reservoir management approach enables the petroleum industry to understand possible ranges of reservoir parameters and optimize production strategy accordingly. Closed-loop reservoir management can also be used to quantify uncertainty in reservoir parameters and take into account during reservoir management process accordingly. An ensemble of reservoir realizations can be incorporated in the workflow to probabilistically forecast production and an optimum production strategy for the overall ensemble can be obtained using robust optimization concepts,. However, robust optimization involves optimizing every realization which requires significant computational cost. Thus, careful consideration is required of the trade-off between the number of models optimized and the computational cost.

This thesis aims to investigate the benefit of optimizing production strategy with different ensemble sizes. Two-phase reservoir modeling of waterflooding is used in this study. Markov Chain Monte Carlo (MCMC) is used in the history matching process to investigate probability distributions of uncertain reservoir parameters. The Minimax approach which aims to maximize spread in input uncertainty space will be used in selecting representative models for different ensemble sizes. Simultaneous Perturbation Stochastic Approximation (SPSA) is applied to each ensemble to optimize production

strategy. The study compared the resulting NPVs using optimized production strategies from different ensemble sizes.

Results show that increasing ensemble size leads to a better development strategy. However, the incremental benefit decreases with increasing ensemble size. The study indicated that the development strategy that is based on multiple realizations is better than development strategy that was developed based on single realization even though the multiple realizations case did not include all possible realizations. The study also demonstrates a systematic methodology for investigating the benefit of using multiple models for optimization vs. a single realization.

DEDICATION

This thesis is dedicated to my family and friends.

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Prof. McVay and Dr. Gildin, and my committee members, Prof. Blasingame, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues and the department faculty and staff for making my time at Texas A&M University a great experience. I also want to extend my gratitude to the Chevron Thailand Exploration and Production, my employer, who provided me financial support throughout my graduate program.

Finally, thanks to my family and friends for their encouragement, patience and love.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION AND LITERATURE REVIEW

1.1 Introduction

In order to maximize reservoir potential, understanding the ranges of reservoir parameters and optimizing development plan are key components. In 2009, Jansen et al. illustrated the concept of the closed-loop reservoir management (CLRM). CLRM is a process that enables us to gain knowledge of reservoir parameters and optimizing development plan in a structured workflow. CLRM has gained growing attention from the petroleum industry recently.

In the conventional data assimilation process, the most probable reservoir realization is obtained during the history matching process. There are numerous studies of methods to determine the most probable reservoir realization, such as maximum-likelihood estimation (MLE) or maximum-a-posterior estimation (MAP) (Rotondi et al. (2006) and Bi et al. (2000)). Since history matching is an ill-posed problem in which different sets of reservoir parameters can reproduce the same set of observed data, one single reservoir realization cannot quantify the reservoir uncertainty. The complex nature of flow behavior inside the reservoir further complicates the history matching process. Brashear et al. (2001) have shown that failing to properly address reservoir uncertainty can lead to suboptimal development plans and inability to maximize reservoir potential. Thus, uncertainty quantification is necessary to enhance reservoir production and economic gain.

Due to highly non-linear behavior of optimization problems in the petroleum industry, several optimization schemes have been investigated (Bieker and Johansen 2007). Uncertainty in reservoir parameters increases complexity of the problem. In the area of optimization under uncertainty, robust optimization (optimizing all the possible realizations) has been identified as having potential to increase ultimate recovery (Van Essen et al. (2009)). However, optimizing a large number of realizations incurs significant computational cost. Thus, typically, only a limited number of reservoir realizations can be optimized. In order to effectively select a limited number of realizations, wise model selection and ranking processes are necessary. Even though estimating computational costs that result from selecting additional realizations for optimization is straightforward, the benefit of incorporating multiple realizations into optimization process is not clearly understood.

1.2 Background

1.2.1 Uncertainty Quantification

The topic of uncertainty assessment has been investigated for decades. Capen (1976) found that people were commonly underestimating uncertainty and suggested that better understanding of uncertainties would have a significant impact on project success. Welsh et al. (2005) published results of a study based on a survey designed to address a number of well-known biases. The study concluded that risk training can offer some advantages in bias-prone situations. However, for oil and gas industry, the experience in the industry offered little to none reduction in bias susceptibility.

Several studies have identified that underestimation of uncertainty can be a source of lower-than-expected returns. Brashear et al. (2001) addressed the issue of low returns in the oil and gas industry during the 1990s with an example of seven simple projects. The study illustrated that the conventional deterministic project selection process caused the underestimation of risk, overestimation of expected value, and misallocation of capital by selecting projects with unnecessary uncompensated risk, which ultimately led to low industry returns. Incorporating risk into the selection process by using full ranges of uncertain costs and reservoir properties, project risk analysis and the use of portfolio optimization can provide more realistic expected values, better understanding of risk and its mitigation, more optimal capital allocation and ultimately improved operational and financial performance.

Even though the concept of uncertainty quantification has been addressed for decades, there is still room for improvement. Bickel and Bratvold (2007) presented the decision-focused uncertainty quantification framework along with results of a survey conducted to determine the status of uncertainty quantification and decision analysis. The response was that uncertainty quantification was limited by lack of time and indicated that the uncertainty quantification and decision analysis process needs to improve in terms of speed and consistency. Hdadou and McVay (2014) published the results of a study to quantify the value of assessing uncertainty and proposed a new framework for quantifying monetary impact due to overconfidence and optimism on portfolio performance. For high risk tolerance relative to portfolio values, moderate overconfidence and moderate optimism will result in an expected disappointment of

about 50% of estimated portfolio value. The study indicated that reducing overconfidence, which prevents consideration of all possible outcomes, will correct other bias, including directional bias. Gonzalez et al. (2013) developed a relational database utilizing the Brier score for tracking probabilistic assessments, applying calibration to improve the probabilistic forecast over time. The examples in their work consist of both petroleum industry and non-petroleum industry problems.

1.2.2 History Matching

Reservoir simulation has been proven to be a valuable tool in reservoir management. History matching is needed to ensure reliable reservoir simulation models. History matching is an inverse problem where the observed data are used for estimating uncertain reservoir parameters.

During the past decades, there has been significant progress in generating reservoir realizations that can match observed production data. Oliver and Chen (2011) published a review on recent progress in history matching. They attributed the progress made in the last decades to increase in computational power and the adoption of geostatistics and Monte Carlo methods. The main components in the history matching process usually consist of (1) parameterization of the uncertain parameters (i.e., zonation), (2) production data and an objective function, (3) algorithms for history matching, and (4) uncertainty quantification methods.

Since a reservoir simulation model typically contains large amount of grid cells, the number of independent variables is significantly higher than the number of observed data points. Thus, it is advantageous to reparameterize the history matching problem to

be in a lower-dimensional space. One of the most widely-used parameterization methods is zonation. However, sub-optimal zonation can easily lead to discontinuous reservoir parameters and violation of reservoir geology. To avoid this problem, several authors have proposed reparameterization algorithms based on prior knowledge or based on data sensitivity (Jafarpour and McLaughlin 2007).

Two key features of production data that differentiate history matching from other inverse problems are observed data that are available at the well locations and non-linear and non-local relationships between model variables and observed data. In general, history matching aims to find the model parameters that minimize the squared error of mismatch between observed data and simulation data. However, having only minimization of data mismatch in the objective function term may lead to reservoir realizations that are significantly different than prior knowledge. Several authors have introduced a Bayesian framework into the history matching process (Rotondi et al. (2006) and Bi et al. (2000)). The objective function in a Bayesian framework will penalize any deviation from prior knowledge of uncertain parameters, which will lead to reservoir realizations that take into account both prior information and observed data.

Even though the forms of objective functions in history matching are pretty harmonious, the algorithms used in history matching vary significantly. History matching can be considered as optimization problem. During the early days of history matching, the process was done manually. Manual history matching relies mainly on good engineering judgment and experience. Users typically implement local regional multipliers around certain wells that experience significant data mismatch. This leads to

loss of geological realism and limited prediction power. As geological realization become more complex, performing manual history matching become more subjective. Several types of algorithms are used in history matching process to automate the process. Some of the widely used methods are genetic algorithms (see Sec. 1.2.4.2) and Markov Chain Monte Carlo (see Sec. 1.2.2.2). Another key benefit of assisted history matching is the capability to generate multiple history-matched models. Since uncertainty in reservoir management is mainly due to limited knowledge in reservoir parameters, using multiple history-matched models can improve ability to explore the uncertainty space. Schaaf et al. (2008) proposed a workflow to reduce reservoir uncertainty using multiple history matched model that can be used to provide reasonably reliable production forecasts. There are three main steps in the proposed workflow: (1) perform experimental design to the whole range of uncertain parameters to identify heavy hitters, (2) generate multiple history matched models using assisted history matching within a Bayesian framework and (3) construct three proxy models of the production forecast through the use of experimental design technique: one for mean of the production forecast, the two others reflecting spread around the mean. Osterloh (2008) demonstrated a method for assisted probabilistic history matching and probabilistic forecast. The proposed method involved using experimental design, response surface modeling, and multiple response optimizations. The desirability function is used for multiple response optimizations. The cumulative production at every two-year interval becomes the target response for multiple response optimizations. The author explained that another useful application of multiple response optimizations is to

select the kth target percentile model that corresponds to the kth target percentile of all the responses.

1.2.2.1 Bayesian Framework

The Bayesian framework is a concept widely used in statistics and probability theory. It provides a way to systematically update the probability of an event with new information. The initial knowledge we have on the reservoir parameters, namely p(m) is called prior knowledge. The probability of the observed data given reservoir parameters namely p(d|m) is called the likelihood probability. The probability of the reservoir parameters given the observed data is called the posterior probability. The Bayes theory is shown in Eq. 1.

$$p(m|d) = \frac{p(d|m)p(m)}{p(d)} \tag{1}$$

where:

p(m|d) is posterior probability

p(d|m) is likelihood probability

p(m) is prior probability

p(d) is probability of observed data

The posterior probability namely p(m|d) contains information on uncertain parameters that take into account both initial knowledge and observed information. The posterior distribution can be used for uncertainty quantification and reservoir management.

1.2.2.2 Markov Chain Monte Carlo (MCMC)

MCMC is a class of algorithms which intend to sample from a target distribution. The algorithm is based on a Markov process that can be described as a system under transition such that the next state depends only on the current state. One of the widely used classes of MCMC is Metropolis-Hasting algorithms. The name originated from papers presented by Metropolis (1953) and Hastings (1970). The MCMC Metropolis-Hasting algorithm is mainly used for obtaining a sequence of random samples from a probability distribution that is difficult to sample from directly. This method can be used for generating a histogram or calculating an expected value of the target distribution. The important advantage of MCMC is that it does not require knowledge of normalizing constant of target distribution. However, the choices of prior, initial guess, and proposal distribution have significant impacts on the burn-in period and the number of iterations required to reach stationarity. For high-dimensional problems, perturbing all parameters at the same time (global perturbation) will lead to low acceptance rates and cause the chain to progress very slowly (Liu and Oliver (2003)).

Burn-in period is the early period of the chain in which the chain is in transition from the initial point to stationarity. Several approaches have been developed to identify the burn-in period. The most widely-used approach is a time-series plot. Another widely used method proposed by Geweke (1992) to check the chain stationarity is to split the chain into two groups after discarding the burn-in period. If the chain reaches stationarity, the mean of the two groups should be the same.

MCMC has been used in conjunction with reservoir simulation for assisted history matching. In conventional MCMC operation, every iteration required a separate reservoir simulation. Thus, applying MCMC to quantify reservoir uncertainty is very computationally expensive. Several modifications were proposed to improve computational efficiency of the MCMC process. The improvements can be grouped into two main areas: (1) replacing the full-field reservoir model with another method; i.e., using experimental design to create a response surface and using it as a proxy for reservoir simulation, and (2) utilize better proposal distribution that is able to increase acceptance ratio. Holmes et al. (2007) proposed a continuous reservoir simulation process that runs through the life of the reservoir while incorporating real time production and pressure information. This method allowed for more runs which enable better uncertainty quantification in production forecasts. The concept was tested on one synthetic reservoir and one field case. The test indicated that this concept is feasible. Liu and McVay (2010) applied the concept proposed by Holmes et al. (2007) in conjunction with the concept of Markov Chain Monte Carlo (MCMC) for exploration of the parameter space to quantify uncertainty in production forecasts. The approach delivers probabilistic production forecasts that narrow with time and provides mechanisms for uncertainty estimation. Alpak et al. (2009) proposed a stochastic history-matching framework that combines the concept of Design of Experiments (DoE) and MCMC. The proposed framework provides ranges for high-impact parameters and multiple history-matched models. However, the proposed framework did not incorporate the production optimization process.

1.2.3 Model Selection

With the improvement in geostatistics and history matching methodologies, hundreds or even thousands of reservoir realizations can easily be generated; thus, reservoir management is moving toward an ensemble-based approach. However, it is very expensive computationally to explore reservoir management strategies using all the realizations. Thus, there is a need to select only models that are statistically representative to be used during the decision making and planning process.

The conventional model selection approach is to select a few representative models at certain percentiles (i.e., 10th, 50th, 90th percentile of original oil in place) of certain reservoir parameter (Deutsch and Srinivasan (1996) and Odai and Ogbe (2011)). The downside of this approach is that the selected models may not be able to capture the uncertainty of other reservoir parameters.

1.2.3.1 Minimax Model-Selection Method

Chen et al. (2013) proposed a new model-selection approach called the "Minimax approach" to select a few models from a large ensemble of models, while maximizing the difference of the models in the input uncertainty space. This approach can match target percentiles of multiple output responses.

The goal of the Minimax model-selection approach is to select the ensemble in a way that not only matches desired target percentiles but also yields an ensemble of models that are maximally different in the input space. This approach requires to solve two combinatorial optimization problems simultaneously (Eq. (2) and Eq. (3)):

$$\min_{X^1,\dots,X^P \in \blacksquare} \sum_{j=1}^{P}\left\{\max_k \min_l \left|\frac{y_k^j - \bar{y}_k^l}{\bar{y}_k^l}\right|\right\} \qquad ,\forall k = 1,\dots M; l = 1,\dots,P \tag{2}$$

$$\max_{X^1,\dots,X^P \in \blacksquare}\{\min_{j,l;j\neq l}(\min_i |x_i^j - x_i^l|)\}, \quad \forall j,l = 1,\dots P; l = 1,\dots,N \tag{3}$$

Subject to:

$$\text{unique}\left(\left[\text{prctile}(y_k^1),\dots,\text{prctile}(y_k^p)\right]\right) = P, \forall k = 1,\dots M \tag{4}$$

where:

M is number or output parameter.

P is number of representative model

y is vector of output parameter

x is vector of input parameter

$\blacksquare = c_1,\dots,c_k$ is set of large but finite model

$\hat{\blacksquare} = c_1,\dots,c_p$ is set of statistically representative model

The objective of Eq. (2) is to minimize the distance of the selected model from the targeted percentile while Eq. (3) aims to find set of models that are maximally different in the input space. Constrain stated in Eq. (4) is that every selected model is unique to each other's.

1.2.4 Optimization Techniques

Multiple optimization techniques have been used in the petroleum industry to assist in the reservoir management process. Optimization algorithms typically start with some initial guess as a solution and the initial solution is then updated iteratively to improve the value of an objective function. Optimization techniques that are widely used

can be divided into two main groups: (1) gradient-based algorithms and (2) evolutionary algorithms.

Gradient-based optimization algorithms refer to a class of optimization that relies on the derivative of the objective function to move from an existing solution to the optimal solution. One of the gradient-based optimization methods that have been gaining industry attraction is the Simultaneous Perturbation Stochastic Approximation (SPSA) (Spall 1998). On the other hand, evolutionary algorithms are inspired by biological evolution. One of the most commonly used evolutionary algorithms is the genetic algorithm (GA) (Bieker and Johansen 2007).

1.2.4.1 Simultaneous Perturbation Stochastic Approximation (SPSA)

Because the gradient-based method uses gradient information to guide the movement from current solution to the next, obtaining gradient information of the objective function is the main step in gradient-based algorithms. The typical gradient-based optimization method for minimization problems can be written in the following form:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k) \tag{5}$$

where:

$\hat{g}_k(\hat{\theta}_k)$ is approximated gradient

$\hat{\theta}_k$ is vector of solution at current iteration

$\hat{\theta}_{k+1}$ is vector of solution at next iteration

a_k is step size

In some situations the gradient information is not available and can be obtained by perturbation. An early prototype gradient approximation is the Kiefer-Wolfowitz algorithm (FDSA). FDSA uses a finite-difference approximation to obtain the gradient of the objective function with respect to each individual element, which makes it easy to implement. However, the main drawback of FDSA is that the method is very computational expensive. For double-sided FDSA, the number of measurements (flow simulations) required for one iteration is twice the number of control variables. Thus, in large systems, FDSA becomes too hard to implement.

SPSA algorithm takes a slightly different approach in approximating gradients. Instead of perturbing one control variable at a time, SPSA randomly perturbs all parameters to obtain two measurements. Then, each component of the gradient is calculated based on the ratio of the individual components in the perturbation vector and the difference of the two measurements. The expected search direction generated from SPSA is the steepest descent generation. Chin (1997) conducted a comparative study of stochastic approximation algorithms, which showed the superiority of the SPSA algorithm. The SPSA algorithm can be expressed as the set of equations below:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - a_k \hat{g}_k(\hat{\theta}_k) \tag{6}$$

$$\hat{g}_k(\hat{\theta}_k) = \frac{y(\hat{\theta}_k + C_k \Delta_k) - y(\hat{\theta}_k + C_k \Delta_k)}{2 C_k \Delta_{ki}} \tag{7}$$

$$a_k = \frac{a}{(k+A)^\alpha} \tag{8}$$

$$C_k = \frac{c}{k^\gamma} \tag{9}$$

Eq. (6) indicates how the solution moves from current iteration to the next iteration. Gradient is approximated by using Eq. (7). Eq. (8) and Eq. (9) are used for calculating the step size and perturbation size respectively.

The steps for implementing SPSA are relatively easy to code. Spall (1998) summarized the key steps in his paper and included sample MATLAB codes. He also proposed guidance for choosing each of the coefficients effectively: $\alpha = 1$ and $\gamma = \frac{1}{6}$ is asymptotically optimum but setting $\alpha = 0.602$ and $\gamma = 0.101$ is effective in practice. Coefficient $'A'$ is suggested to be five to ten percent of the number of iterations. Coefficient $'c'$ is recommended to be set equal to the standard deviation of noise in the objective function. Coefficient $'a'$ is recommended to be set in a way that the product between a_k and g_k is equal to the smallest desired step size. Δ_k is user-specified random perturbation vector. The Bernoulli distribution with probability of ½ for ± 1 outcome is recommended for the perturbation vectors. The key benefits of SPSA are significant reduction in computation time required for approximating gradients, especially in high-dimensional problems, and ease of implementation. Gao et al. (2007) implemented SPSA for automatic history matching.

1.2.4.2 Genetic Algorithm (GA)

The genetic algorithm (GA) can be considered as a search algorithm that is based on the process of natural genetics. Since GA is not a solver algorithm, there is no need for knowledge of the objective function shape. The steps for GA can be summarized as follows: (1) create random initial population, (2) score each member of the current population and scale the values, (3) select member based on their fitness value to be

parents for next generation, (4) produce children from their parents by combining vector of parents (crossover) or randomly change part of parent vector (mutation), (5) carry individuals with the best fitness value in current generation to the next generation (elite children) (6) replace current population with the new population and (7) repeat from step (1) until optimal criteria are met.

The key advantage of GA is that the algorithm can explore a wide parameter space and not be easily trapped in local minima. However, as GA does not use any gradient information to guide to the optimum solution, GA is very computational expensive. In order to improve the slow rate of convergence, gradient-based search methods are used in conjunction with GA to improve the local convergence rate.

1.2.5 Closed-Loop Reservoir Management (CLRM)

Closed-loop reservoir management is a combination of model-based optimization and assisted history matching. Recently, closed-loop reservoir management has gained growing attention from the petroleum industry and in the context of 'i-fields' or 'smart fields.' The key concept of CLRM is to maximize reservoir performance over the life of the reservoir by changing the reservoir management process from a periodic to near-continuous process. Jansen et al. (2009) utilized closed-loop reservoir management as a continuous process for history matching, optimizing NPV as new information is obtained. Gildin et al. (2011) proposed a general workflow to utilize real-time optimal-control strategies for large-scale reservoir models. Pajonk et al. (2011) investigated the potential of increasing oil recovery through the use of smart-well technologies. The ensemble-based hybrid optimization workflow comprising an ensemble Kalman filter

15

and the Covariance Matrix Adaptation Evolution Strategy was proposed. The ensemble Kalman filter was used for history matching and the Covariance Matrix Adaptation Evolution Strategy for production optimization to provide a production strategy that yields the highest expected NPV from the whole ensemble.

1.2.5.1 Robust Optimization (RO)

Dynamic optimization has been recognized by the petroleum industry as having potential to increase ultimate recovery and profitability. However, dynamic optimization often lacks flexibility to incorporate uncertainty in geological parameters. Robust optimization (RO) utilizes multiple sets of geological realizations in the optimization process, which leads to a development plan that is optimum for all realizations. Van Essen et al. (2009) compared robust optimization with nominal optimization and reactive-control strategies and found that robust optimization is superior to the other two strategies. However, the comparison is based on a situation in which no production data is available. Hence, each realization is assumed to have equal probability of being the true model. Alhuthali et al. (2008) proposed a practical approach to determine the optimum production and injection rates under geological uncertainty. Geological uncertainty is handled by using two optimization problems. Objective function of the first optimization relies on combination of expected value and standard deviation while the second objective function focuses on minimizing the worst case scenario.

1.3 Research Objective

This study is motivated primarily by the Van Essen et al. (2009) work on robust optimization. Due to advancements in data assimilation processes, multiple reservoir

realizations can be obtained (Oliver and Chen (2011)). Optimizing the production strategy for every realization is possible but with significant computational cost; thus, decision on how many models to be used during optimization process must be made carefully.

The objective of this study is to investigate the benefit of optimizing production strategy with different ensemble size.

2. METHODOLOGY AND MODEL DESCRIPTION

2.1 Research Workflow

This study will utilize a Bayesian framework and MCMC to assist in data assimilation to investigate the range of each uncertain parameter. Knowledge from data assimilation will be used in model selection for further optimization. Flow simulation used in this study is ECLIPSE. The steps taken in this study are as follows:

1. Generate production profile using a synthetic reservoir model. This model will be used as the "true" reservoir model. The production will be under waterflooding development. Additional noise will be incorporated into the generated production from the "true" reservoir model to mimic noise in field production.

2. Perform data assimilation to obtain distributions of uncertain parameters.

3. Select different ensembles of models with different ensemble sizes to optimize the production strategy. Parameters for the optimization process are production and injection rates.

4. Input the optimum strategy from each ensemble size to all possible realizations to obtain the resulting NPV. The resulting NPV will be used to investigate the benefit from different ensemble size. The resulting production forecast from different ensemble size will also be compared with the true model to investigate the range of production forecast.

2.2 Model Description

The model used in this study is based on the Brugge field case. The Brugge field case was a synthetic model developed as a realistic case study for benchmarking waterflooding history matching and optimization techniques (Peters et al. (2013)). The setup of the Brugge field case is similar to the PUNQ field case, which focuses on uncertainty quantification. The main difference between two cases is that the Brugge case was intended for participants to develop optimal waterflooding plans.

The structure of Brugge field consists of an east to west elongated half-dome with a large boundary fault on the north side of the reservoir. There is one internal fault on the north side of the reservoir (Fig. 1). Field areal extent is roughly 10×3 km. Reservoir properties and thickness are typical of North Sea Brent-type fields.

Brugge field contains ten years of production history with ten water injectors and twenty producers. All injectors were placed down dip of the reservoir in the peripheral manner. Table 1 summarizes general information on the Brugge simulation model.
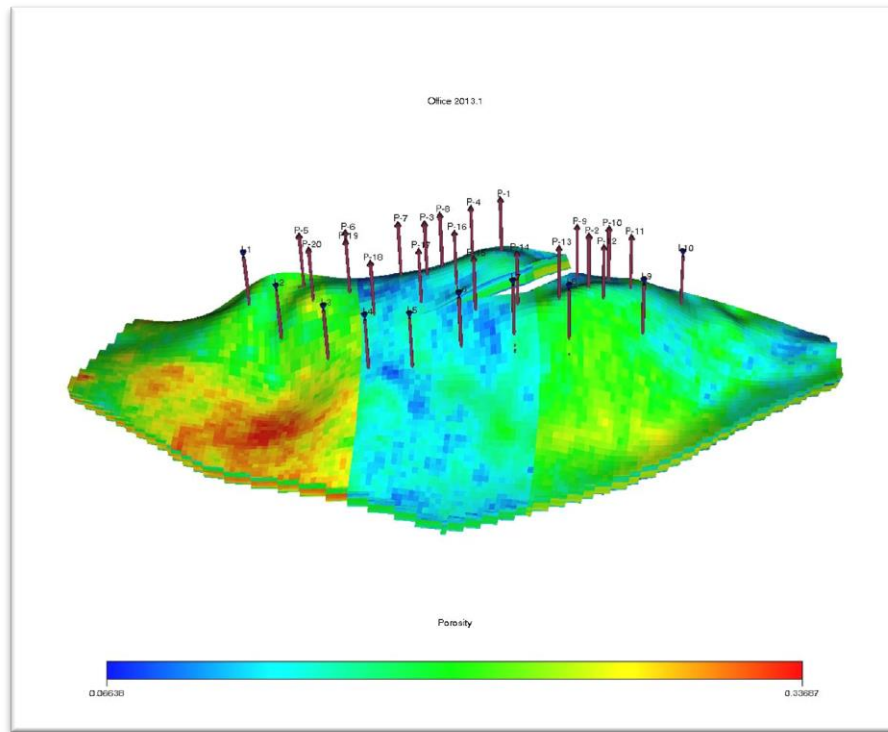
Fig. 1—Brugge field model

Table 1—General information on Brugge field

Parameter	Value	Unit
Number of grids (DX x DY x DZ)	139 x 48 x 9	cell
Number of faults	1	
Initial pressure	2480	psi
OWC	5505	ft.
Relative permeability correlation	Corey	
Rock compressibility	3.50E-06	psi^{-1}
PVT	Dead oil	
Fluid type	Oil, Water	
Oil viscosity	1.294	cp
Oil density	56	lbm/ft^3
Oil compressibility	9.26E-06	psi^{-1}
Water viscosity	0.32	cp
Water density	62.6	lbm/ft^3
Water compressibility	3.00E-06	psi^{-1}

2.3 Truth Case Definition and Parameterization

The truth case reservoir used in this study is based on the Brugge case with some modifications. The reservoir is arbitrary divided into three regions (Fig. 2).
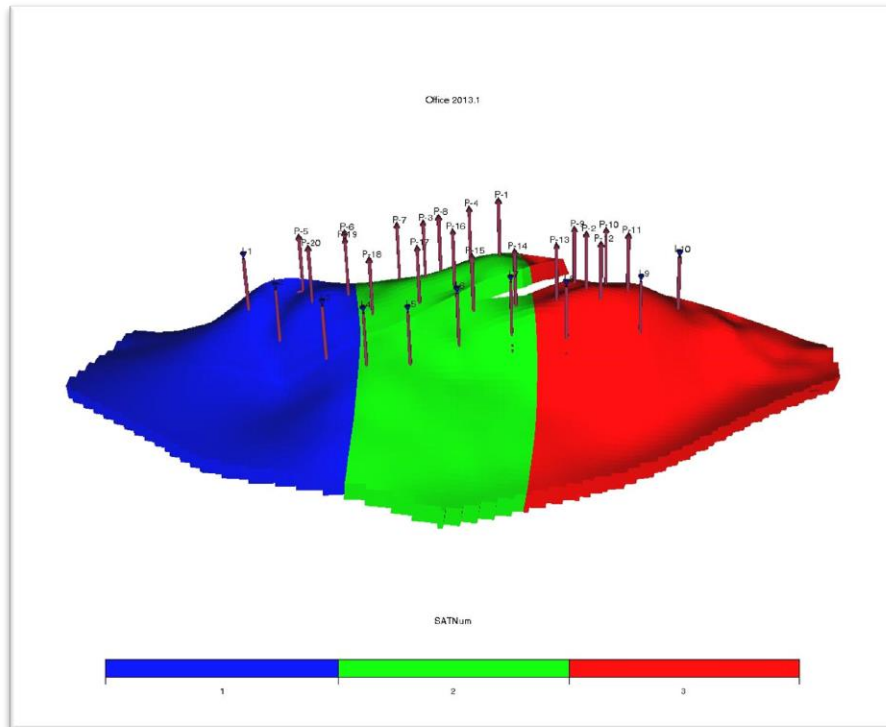
Fig. 2—Reservoir region

Four uncertain reservoir parameters in each of the three regions are random to create the truth case. Uncertain parameters are ratio of horizontal permeability to vertical permeability, horizontal permeability multiplier, end-point water relative permeability at water saturation of 0.75, and porosity multiplier. **Table 2** summarizes the reservoir parameters of the truth case.

Table 2—Truth case reservoir parameters

Region	k_v/k_h Ratio	k_h Multiplier	k_{rw} end point	Porosity Multiplier
1	0.07	0.1	0.35	1.25
2	0.13	5	0.54	0.78
3	0.15	3	0.42	1.04

2.4 Observed Data and Prior Distribution

The parameters in **Table 2** were applied to the Brugge field case. Then, oil production rate and water injection rate were controlled according to Brugge historical data to obtain production profiles. Since field production data typically contains some noise due to measurement errors, additional noise was added to the profile obtained from flow simulation. Noise added to the observed data in this study had normal distributions with mean and standard deviations shown in **Table 3**. Frequencies of each of the parameters are also summarized in **Table 3**. Production profiles after incorporating noise will be used as observed data in history matching.

Table 3—Observed data noise distributions

Parameter	Frequency	Noise mean	Noise standard deviation
Oil production rate	Monthly	0	1% of true value
Well flowing bottomhole pressure	Monthly	0	15 psi
Water cut	Monthly	0	3% of true value
Water injection rate	Monthly	0	3% of true value

Prior distributions in this study were of the truncated Gaussian distribution form.

Information on parameter distributions is summarized in **Table 4**, **Table 5** and **Table 6**.

Table 4—Prior mean

Region	k_v/k_h ratio	k_h multiplier	k_{rw} end point	Porosity multiplier
1	0.1	1.00	0.375	1.00
2	0.1	1.00	0.375	1.00
3	0.1	1.00	0.375	1.00

Table 5—Prior standard deviation

Region	k_v/k_h ratio	k_h multiplier	k_{rw} end point	Porosity multiplier
1	0.02	0.5	0.056	0.15
2	0.02	0.5	0.056	0.15
3	0.02	0.5	0.056	0.15

Table 6—Prior: Parameter range

Parameter	Minimum	Maximum
k_v/k_h ratio	0.01	0.25
k_h multiplier	0.05	10.00
k_{rw} end point	0.20	0.55
Porosity multiplier	0.50	1.50

3. HISTORY MATCHING

3.1 Introduction

In this section, I will explain the MCMC process employed, a sensitivity study performed on the MCMC process, and the resulting posterior distribution that was used in model selection and production optimization. One of the main objectives of history matching is to obtain reliable reservoir models, with uncertainty, to be used in reservoir management. The main idea of uncertainty quantification is to gain an understanding of the distribution of each uncertain parameter.

One of the widely used history matching techniques in uncertainty quantification is Markov Chain Monte Carlo (MCMC). The class of MCMC used in this study is MCMC Metropolis-Hasting algorithm with random walk. Steps for performing history matching in this study are:

1. Generate an initial state of each uncertain parameter.

2. Run a simulation model based on the initial state.

3. Calculate the posterior probability of the initial state.

4. Randomly select the next possible state based on the proposal distribution and current state.

5. Run a simulation model based on the new state.

6. Evaluate probability of the new state and calculate the acceptance probability.

7. If the new state is accepted, the next state will be based on the new state, but if the new state is rejected, the next state will be based on the previous state.

8. Repeat from step 4 until a stationary chain is obtained or the chain reaches the maximum number of iterations.

Choices of initial guesses and proposal distribution significantly impact the convergence rate of the chain. The common choice for initial guesses is based on prior knowledge of each uncertain parameter. Success of the Metropolis-Hasting algorithm depends on not having too low of an acceptance rate. Using small step size will lead to a higher acceptance rate. However, with small step sizes many iterations are needed to explore the whole parameter space in order to converge to a target distribution. Large step size can be used but the acceptance rate will be low, especially in the tail region. The important components of MCMC will be described in detail in the following section.

3.1.1 Proposal Distribution

Choice of proposal distribution plays an important role in the performance of MCMC in history matching problems. The proposal distribution is used for generating a proposed state based on the current state. The proposal distribution used in this study is the random walk scheme, which is the most common and practical option for proposal distributions. In the random walk scheme, the proposed state is based on the current state plus some random variable (Eq. 10). Distribution for $\mathbf{w_j}$ in this study is a normal distribution.

$$m_{i+1} = m_i + w_j \tag{10}$$

where:

m_i is uncertain parameter vector at current state

m_{i+1} is uncertain parameter vector at proposed state

w_j is random variable with distribution independent of the chain

3.1.2 Prior Distribution

In this study the prior distribution is assumed to have a normal distribution, which can be written in the following form:

$$p(m) = c \times exp[-\frac{1}{2}(m_i - m_o)^T C_M^{-1}(m_i - m_o)] \tag{11}$$

where:

c is normalizing constant

m_i is uncertain parameter vector at current state

m_o is uncertain parameter vector as per prior knowledge

C_M is prior covariance matrix

Uncertain parameters in this study have no correlation between each parameter. The prior covariance matrix is a diagonal matrix with variance of the prior for each parameter. Values of prior standard deviation represent levels of uncertainty for uncertain parameters, and these standard deviations are typically large due to large uncertainty in the parameters. In this study, there are twelve uncertain parameters; thus, the uncertain vector dimension is 12×1.

3.1.3 Likelihood Function

The likelihood function represents the probability that the given state will produce the observed data (p(d|m)). As mentioned in section 2.4, the observed data in this study consists of four parameters. Oil production rate and water injection rate will be

used as control parameters for history matching. Well flowing bottomhole pressure and water-cut will be used for calculating the likelihood function. Because there are ten water injectors and twenty oil producers with monthly data frequency, the total data points for the likelihood calculation is 6,000 data points.

The distribution of noise in this study is of the Gaussian distribution form with standard deviation as shown in section 2.4. The likelihood function can be written as:

$$p(d|m) = c \times exp[-\frac{1}{2}(d_{obs} - g(m))^T C_D^{-1}(d_{obs} - g(m))] \tag{12}$$

where:

p(d|m) is likelihood probability

d_{obs} is observed data

g(m) is production profile from flow simulation

C_D is likelihood covariance matrix

The general practice for calculating likelihood probability is to assume that the data mismatch between each point is independent and not correlated to each other. Thus, the likelihood covariance matrix is a diagonal matrix with variance of measurement error for each parameter.

3.1.4 Posterior Distribution

The posterior distribution ($p(m|d)$) is the probability that the state is true given the observed data. The relationship between prior distribution, likelihood function, and posterior distribution is given in section 1.2.2.1. Given the prior probability and likelihood function in Eq.(11) and Eq. (12), the posterior probability can be written as:

$$p(m|d) = c \times exp[-\frac{1}{2}((m - m^0)^T C_M^{-1}(m - m^0)$$
$$+ (d_{obs} - g(m))^T C_D^{-1}(d_{obs} - g(m)))] \tag{13}$$

Eq. (13) can also be written in the following form:

$$p(m|d) = c \times exp[-O(m)] \tag{14}$$

where

$$O(m) = ((m - m^0)^T C_M^{-1}(m - m^0) + (d_{obs} - g(m))^T C_D^{-1}(d_{obs} - g(m)) \tag{15}$$

The term $O(m)$ is typically referred to as the objective function of the posterior distribution, which combines objective functions from the prior and likelihood. The posterior probability takes into account both deviation from prior knowledge and mismatch between observed data and simulation output.

3.1.5 Acceptance Probability

Acceptance probability (α) is the probability that the chain will move from the current state to the proposed state. Hastings (1970) proposed to define acceptance probability in a way that when combined with a transitional kernel, the chain becomes reversible. The acceptance probability can be written as:

$$\alpha = min(1, \frac{p(m_{i+1}|d_{obs})}{p(m_i|d_{obs})}) \tag{16}$$

In summary, the acceptance probability is the ratio of posterior probability between the proposed state ($p(m_{i+1}|d_{obs})$) and the current state ($p(m_i|d_{obs})$). If the proposed state has higher probability than the current state, the proposed stated will be

accepted. If the proposed state has lower probability than the current state, the proposed state may still be accepted. This allows the chain to explore the whole uncertainty space.

3.2 Uncertainty Quantification and Sensitivity Study

In this section, we perform sensitivity experiments on several aspects of history matching using MCMC to understand the performance of the history matching process and to select proper parameters for history matching that will be used for further study. The observed data for history matching is shown in Fig. 3, Fig. 4 and Fig. 5. These data were generated as explained in Section 2.4.
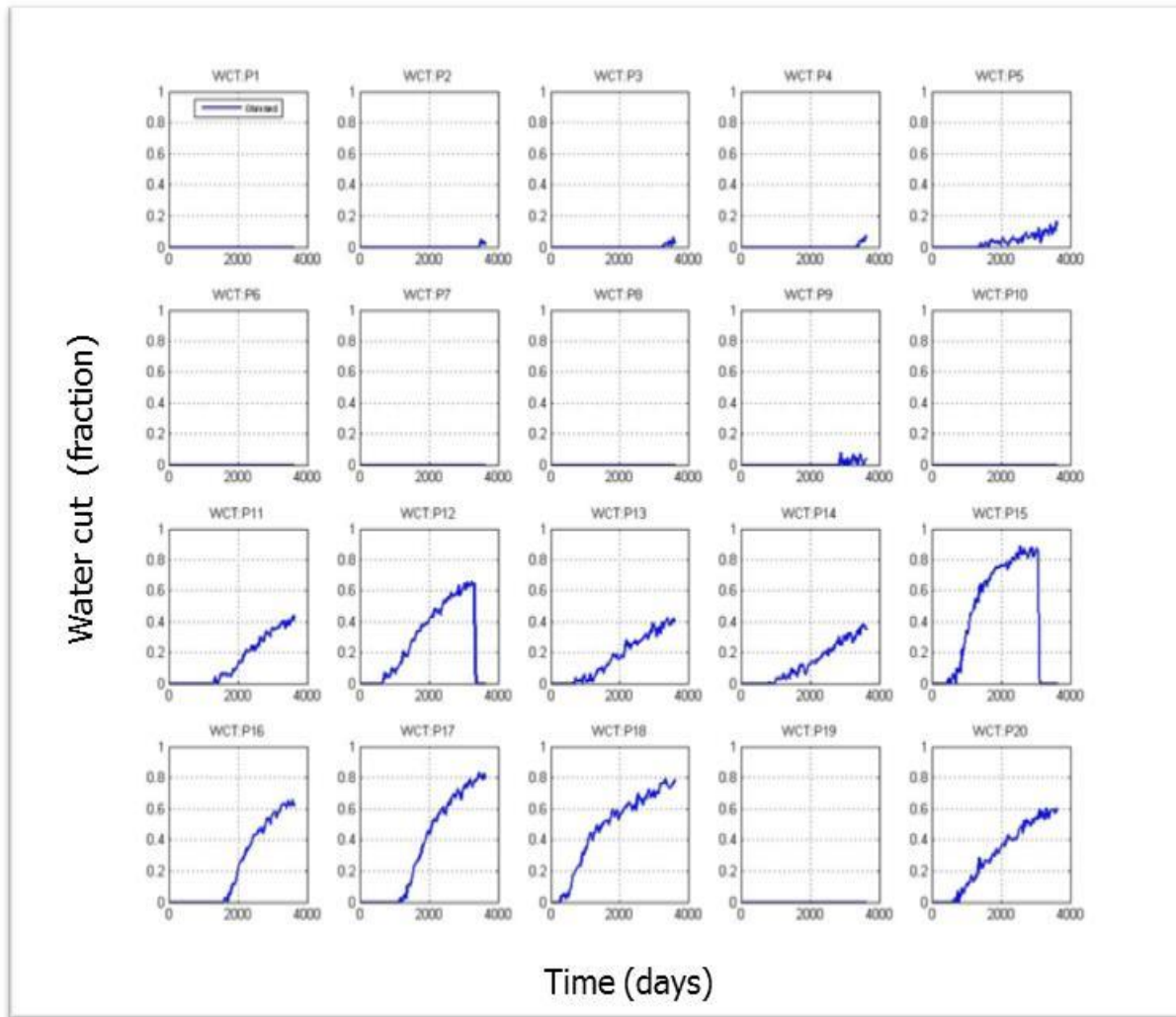
Fig. 3—Producer water cut vs. time
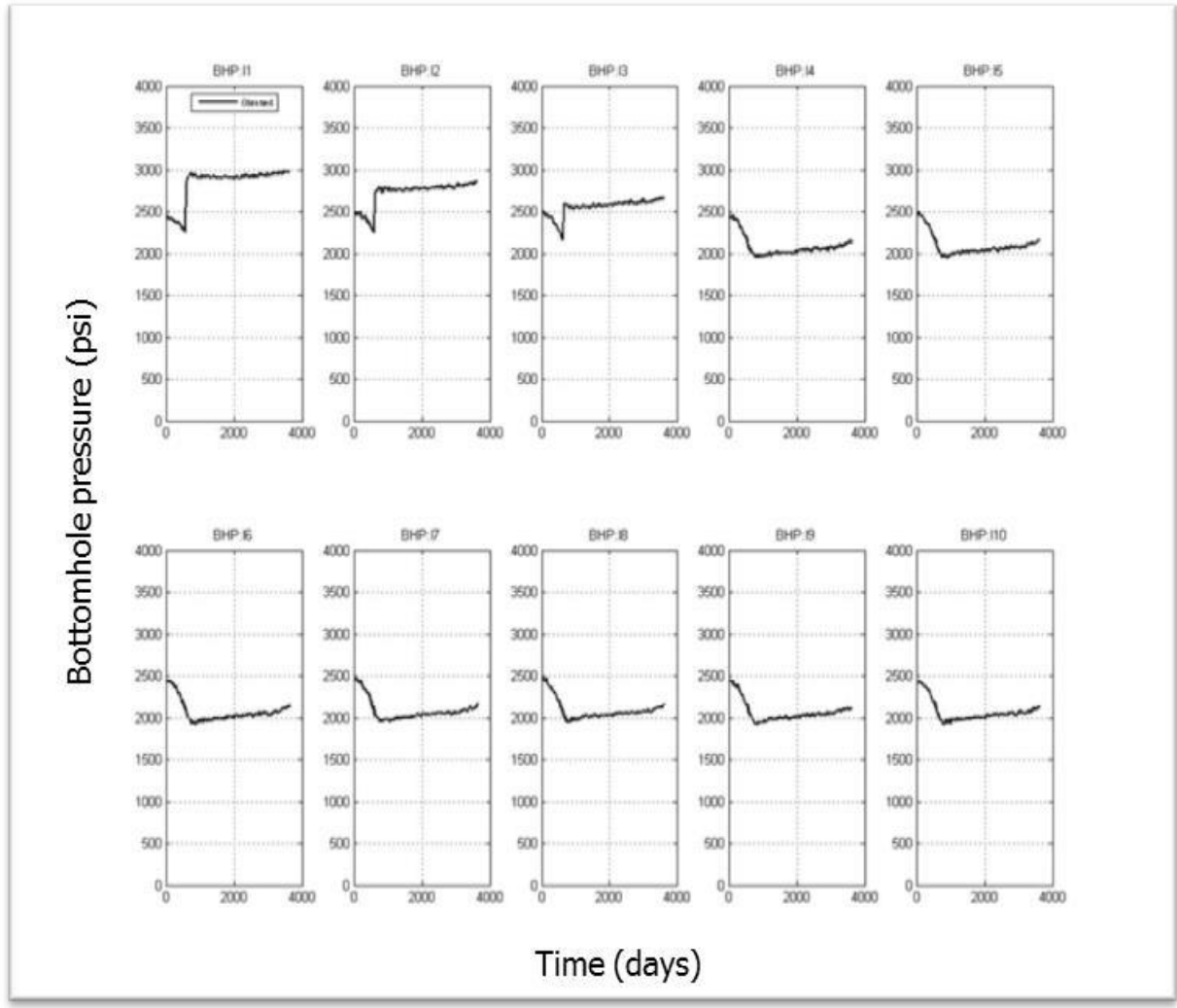
Fig. 4—Injector bottomhole flowing pressure vs. time
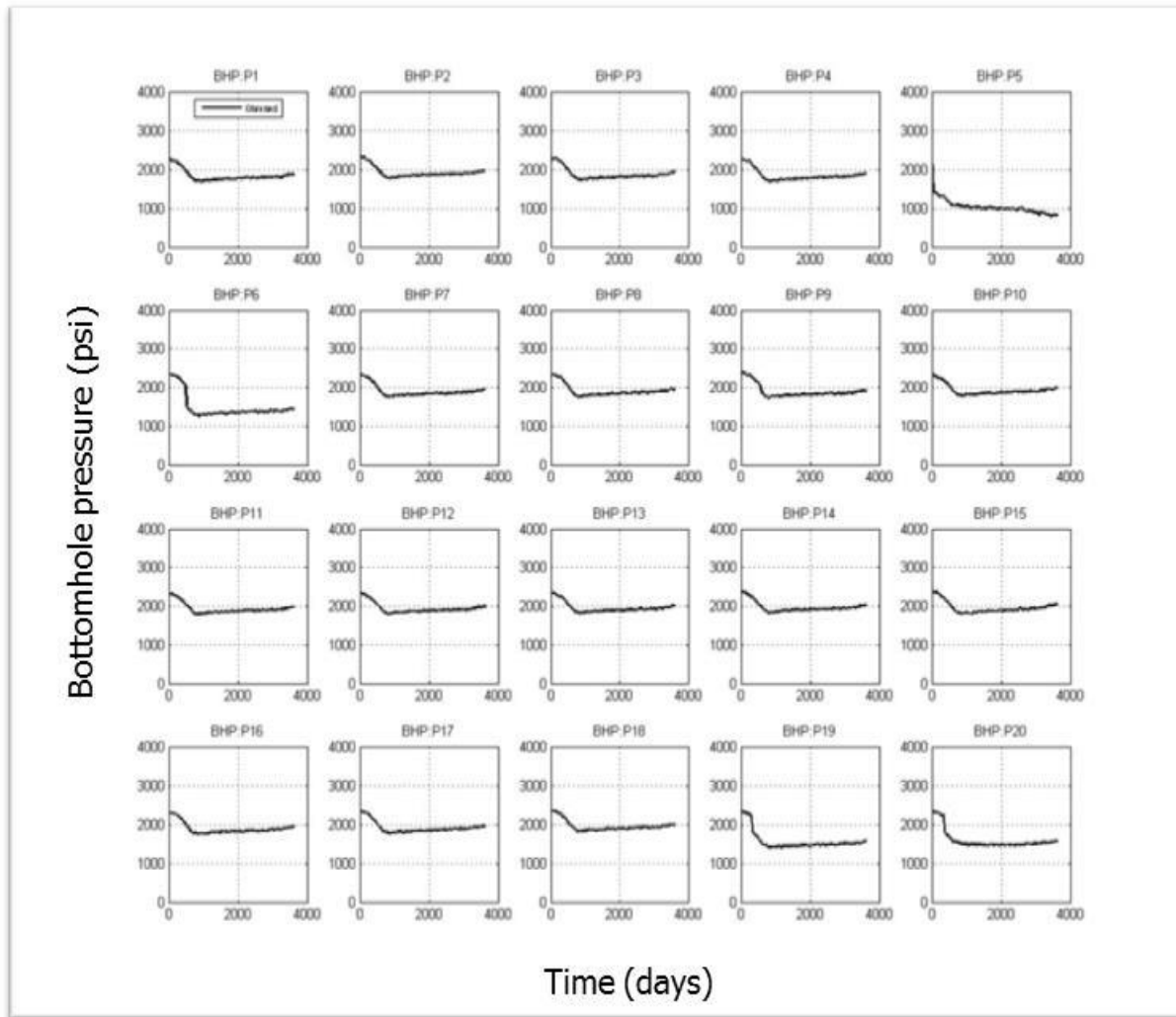
Fig. 5—Producer bottomhole flowing pressure vs. time

33

3.2.1 Effect of Numbers of Perturbation Parameters and Proposal Step Size

In this experiment, I compare the performance of global perturbation (perturbing all uncertain parameters) and local perturbation (perturbing only some uncertain parameters) at different proposal step sizes. Oliver (1997) explained the advantages of local perturbation, including high acceptance rate. The purpose of this experiment is to understand the convergence rate of MCMC at different proposal step sizes and different numbers of variables perturbed.

In local perturbation, four parameters are perturbed during each iteration. Each of the scenarios is run until the chain reaches one thousand iteration. In this study, the burn-in period is approximated based on the iteration that value of objective function start to stabilize. **Table 7** summarizes the scenarios and acceptance rates in this experiment. The relationships between objective function vs. number of iterations for each scenario are shown in Fig. 6 and Fig. 7.

Table 7—Scenarios to investigate perturbation scheme and proposal step size

Perturbation scheme	Proposal step size	Acceptance rate whole chain	Acceptance rate after burn-in Period
Global perturbation	$0.05 *\sigma_{prior}$	8.50%	3.60%
Global perturbation	$0.10 *\sigma_{prior}$	4.80%	2.10%
Global perturbation	$0.25 *\sigma_{prior}$	2.10%	0.90%
Local perturbation	$0.05 *\sigma_{prior}$	30.10%	16.10%
Local perturbation	$0.10 *\sigma_{prior}$	18.10%	11.90%
Local perturbation	$0.25 *\sigma_{prior}$	9.00%	5.30%
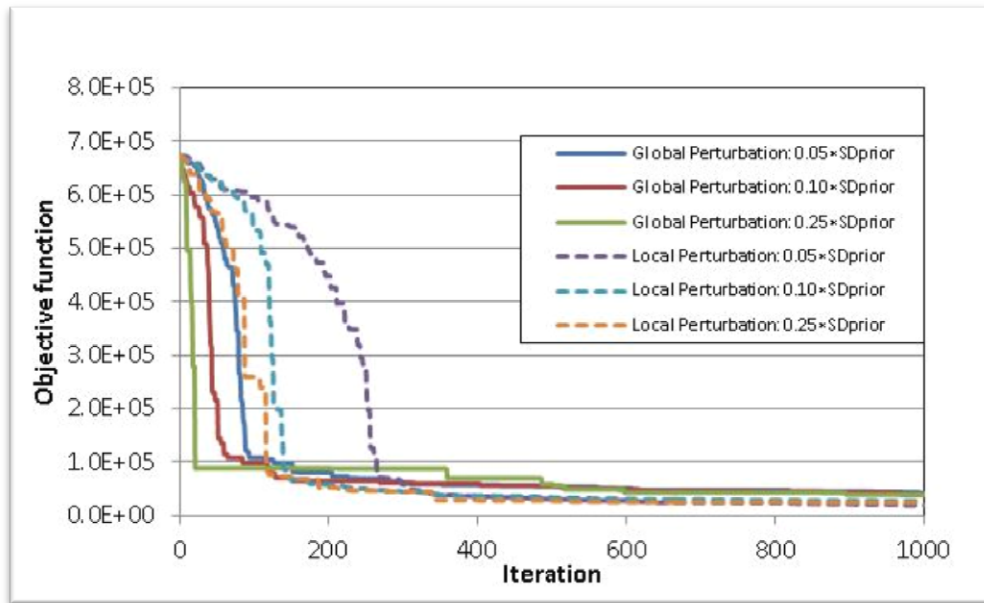
Fig. 6—Effect of Numbers of Perturbation Parameters and Proposal Step Size
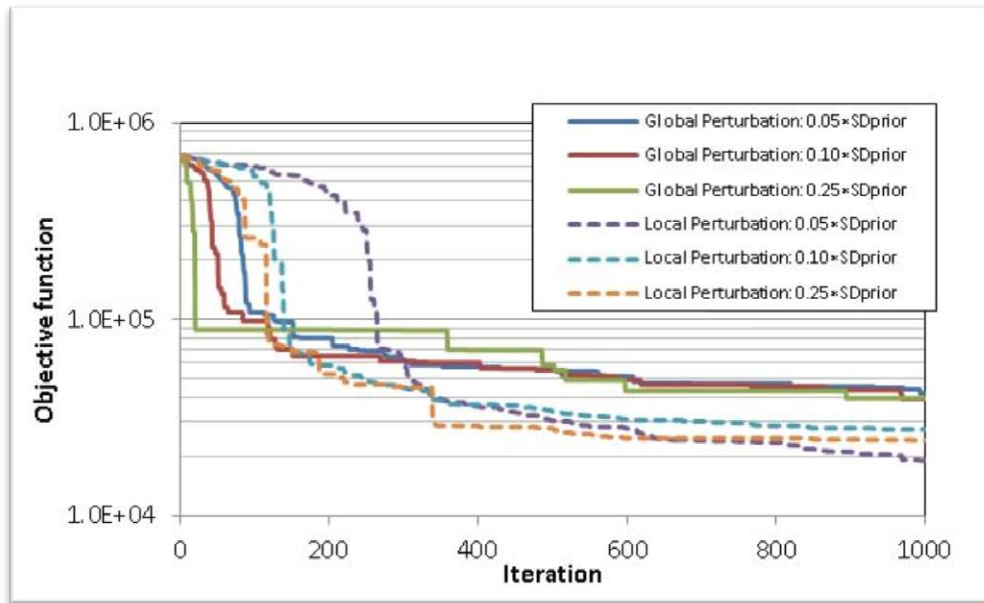
Fig. 7—Effect of Numbers of Perturbation Parameters and Proposal Step Size (Semi-log)

In all cases, global perturbation has significantly lower burn-in periods. This is because global perturbation changes a higher number of parameters between each iteration compared to local perturbation. However, the acceptance rate for global perturbation is significantly lower, especially with larger proposal step sizes. The local perturbation scheme has longer burn-in periods but higher acceptance rates, which lead to better performance when the whole chain is considered (see Fig. 7).

3.2.2 Effect of Prior Choice

In this experiment, we investigate the effect of prior standard deviation (σ_{prior}) on MCMC performance. The value of σ_{prior} has a direct impact on the objective function in the prior term. The value of σ_{prior} relates to our understanding of each uncertain parameter. Prior standard deviations are shown in section 2.4. **Table 8** summarizes settings in this experiment.

Table 8—Parameter setting for investigating effect of $\boldsymbol{\sigma_{prior}}$

σ_{prior}	Proposal Step Size	Perturbation Scheme
0.2*σ_{prior} Basecase	0.05*SD	Local perturbation (4 parameters)
0.5*σ_{prior} Basecase	0.05*SD	Local perturbation (4 parameters)
σ_{prior} Basecase	0.05*SD	Local perturbation (4 parameters)
2*σ_{prior} Basecase	0.05*SD	Local perturbation (4 parameters)
5*σ_{prior} Basecase	0.05*SD	Local perturbation (4 parameters)
10*σ_{prior} Basecase	0.05*SD	Local perturbation (4 parameters)
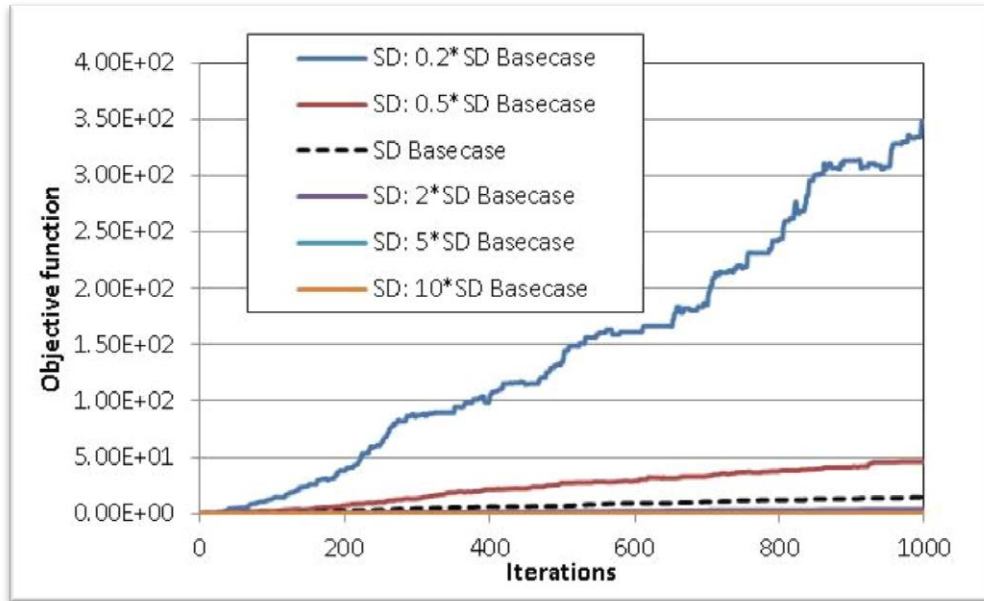
Fig. 8—Prior objective function value vs iteration at different σ_{prior}

Fig. 8 shows that the lower the σ_{prior} value the bigger the prior term. This is because uncertain parameter that differ from prior knowledge get penalize more compare to case with higher σ_{prior}. However, due to significant amount of observed data, the posterior distribution is mostly the result of contribution from the likelihood function. As shown in Fig. 7, the magnitude of the posterior distribution is in the range of 10^5 while objective function of the prior term is only in the magnitude of 10^2. Thus, the knowledge of σ_{prior} value in this study does not have a significant impact on history matching.

3.2.3 Effect of Likelihood Covariance

The objective of this experiment is to investigate the effect of noise standard deviation (σ_{noise}) on the progress of MCMC in this study. The base case of σ_{noise} is presented in section 2.4. Even though we know true σ_{noise} in this study, in some cases it

is difficult to obtain the true σ_{noise}. Thus, understanding the magnitude of impact from knowledge of σ_{noise} value to objective function value and chain acceptance rate is necessary. The value of σ_{noise} for each case and resulting acceptance rate is shown in **Table 9**. Results of this experiment are shown in Fig. 9.

Table 9—σ_{noise} value for experiment 3.2.3 and resulting acceptance rate

σ_{noise} Value	Acceptance rate whole chain	Acceptance rate after burn-in Period
0.2* σ Base case	28.00%	15.00%
0.5*σ Base case	28.00%	14.90%
σ Base case	30.10%	16.10%
2* σ Base case	29.00%	16.40%
5* σ Base case	36.10%	22.20%
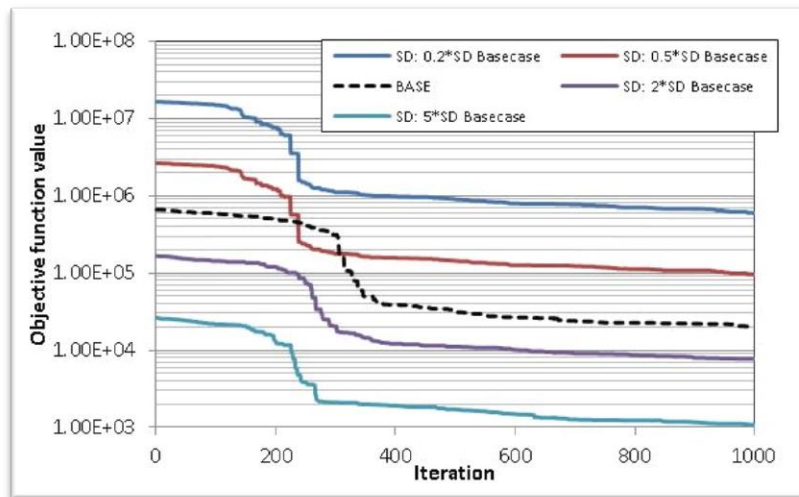
Fig. 9—Posterior objective function value vs iteration at different σ_{noise} (Semi-log)

The experiment shows that the knowledge of σ_{noise} value has a significant impact on objective function of the posterior distribution. This is due to the large amount of observed data in this study, because any change in σ_{noise} will be magnified by the number of available observed data.

3.3 History Matching Case for Optimization

3.3.1 History Matching Quality

In order to understand the distribution of uncertain parameters, we need to obtain stationary MCMC chains. The MCMC local perturbation scheme with perturbation of four parameters was run to three thousand iterations. However, the acceptance rate drops significantly as the chain progresses. This leads to a very small number of accepted models after the burn-in period. Thus, the number of perturbed parameters was reduced to only one parameter at a time. Fig. 10 shows the value of objective function vs. iteration. The acceptance rate when perturb only one parameter is 31%.

By only perturbing one parameter at a time the acceptance rate is improved, which leads to a greater number of accepted realizations. Fig. 11, Fig. 12 and Fig. 13 show plots of observed and simulated water cut and bottomhole flowing pressure at the start of the chain.
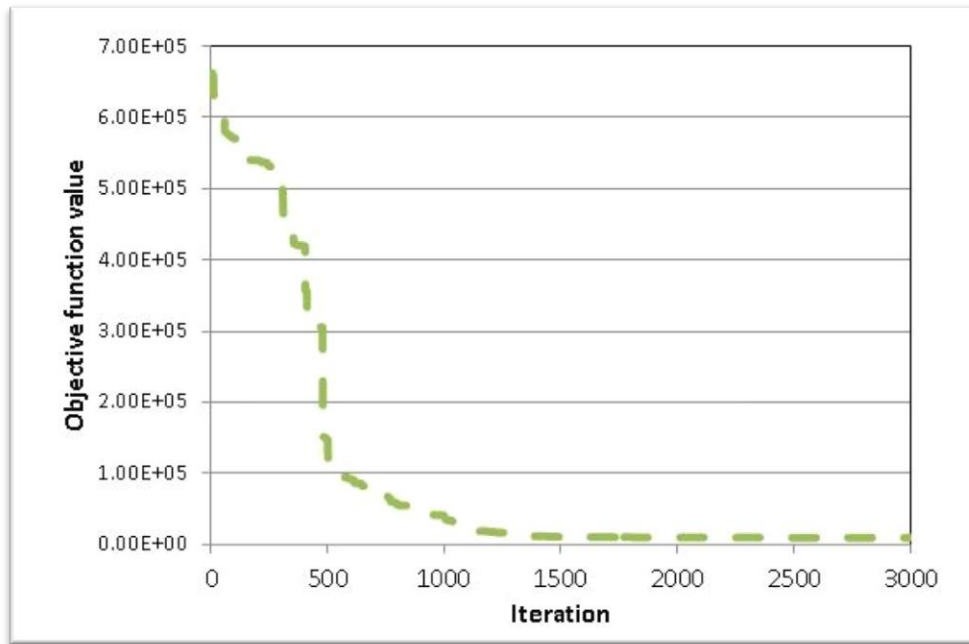
Fig. 10—Objective function vs iteration with one variable perturbed at a time

As shown in Fig. 11, Fig. 12 and Fig. 13, there is significant data mismatch at multiple wells in all three variables. The results of the match at the end of the chain are shown in Fig.14, Fig. 15 and Fig. 16. We observe significantly lower data mismatch.
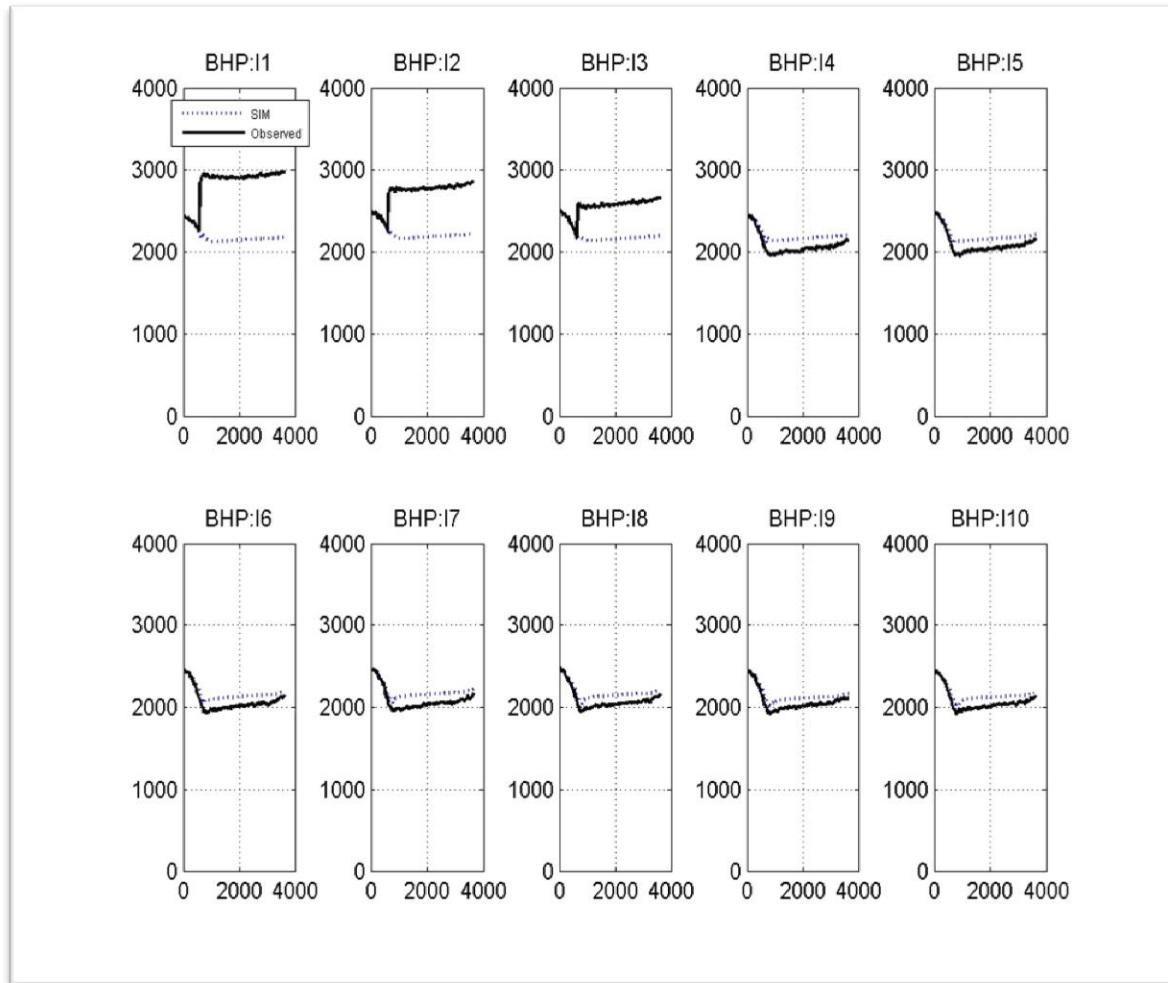
Fig. 11—Observed and simulated bottomhole flowing pressure of water injectors vs. time with prior knowledge.

41
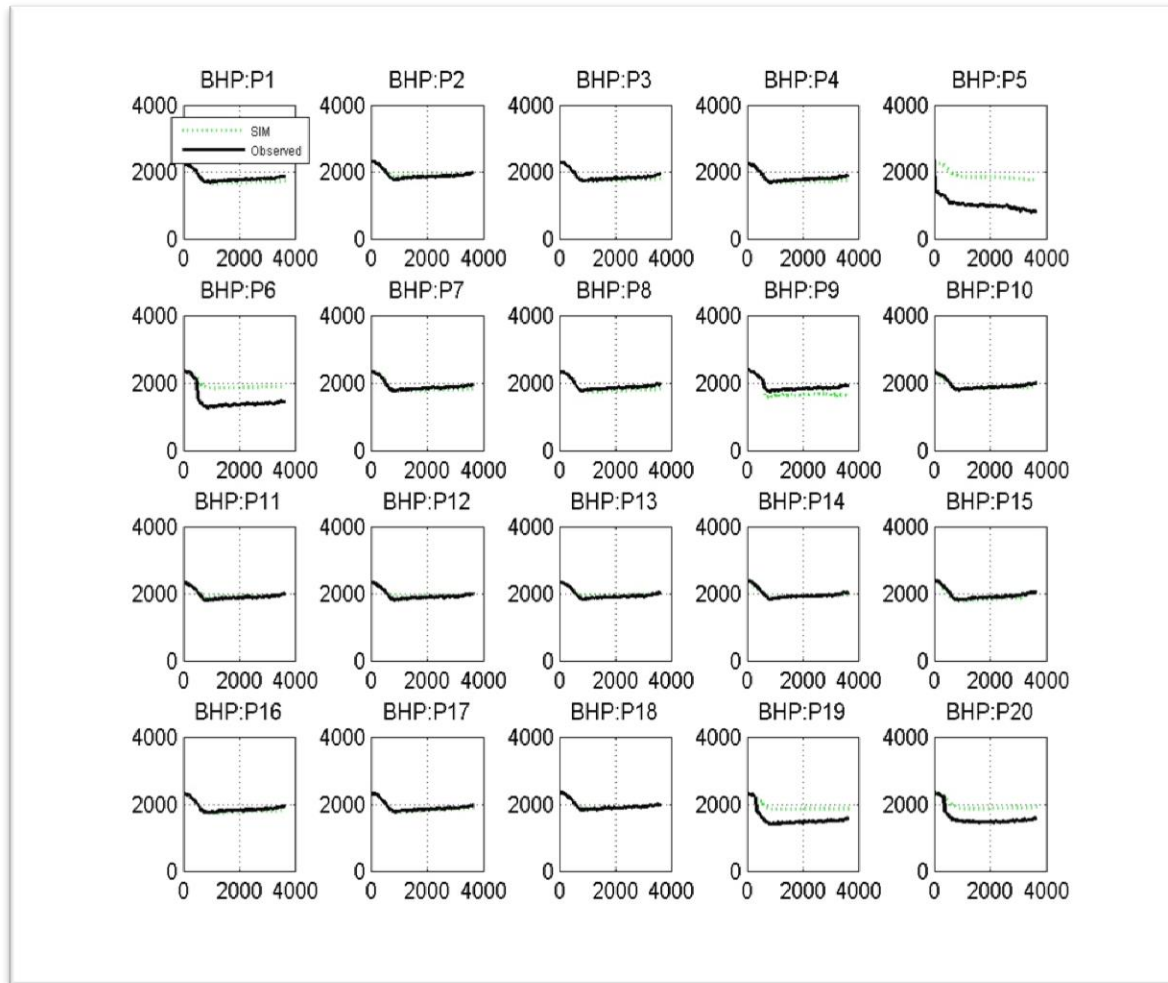
Fig. 12—Observed and simulated bottomhole flowing pressure of producers vs. time with prior knowledge.
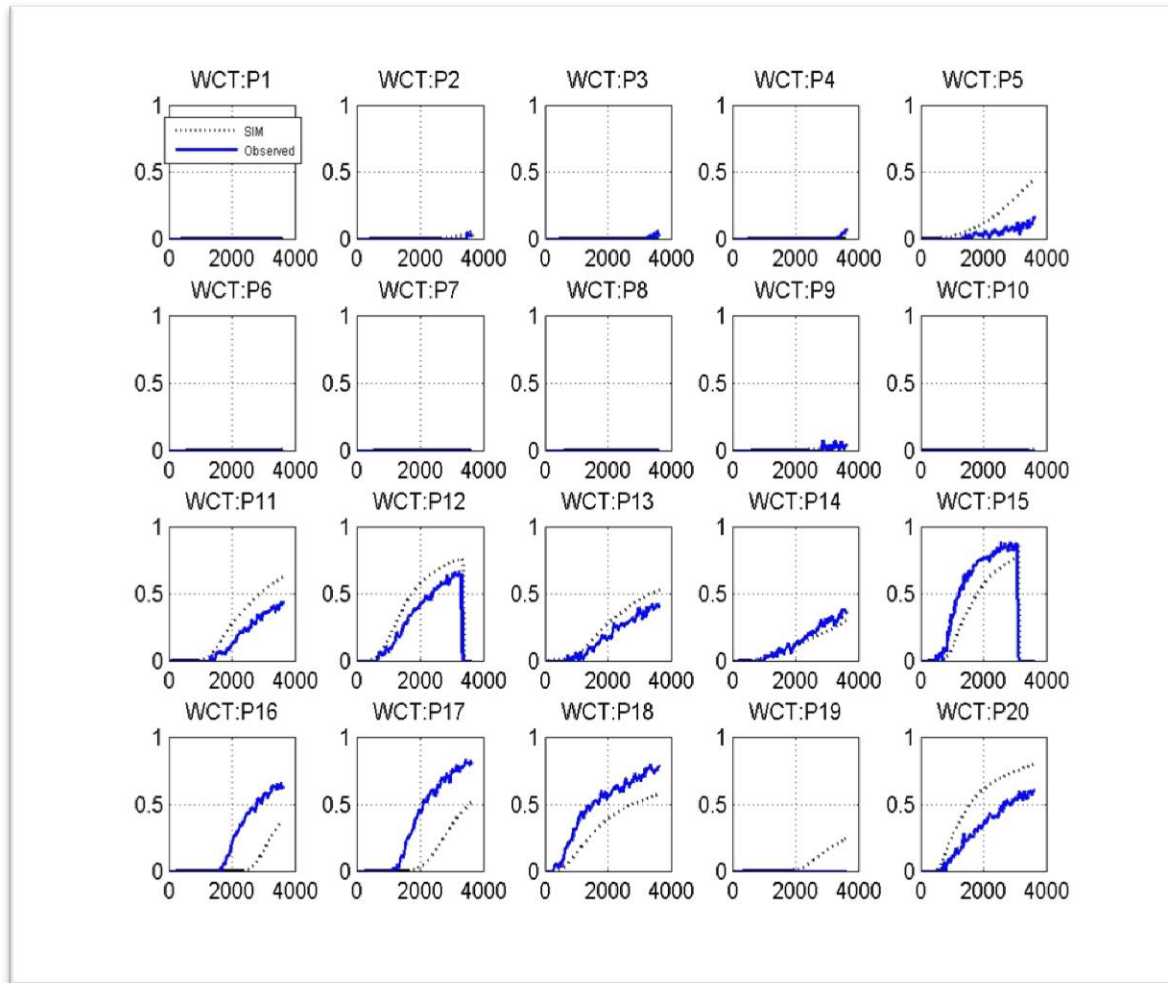
Fig. 13—Observed and simulated water cut of producers vs. time with prior knowledge.
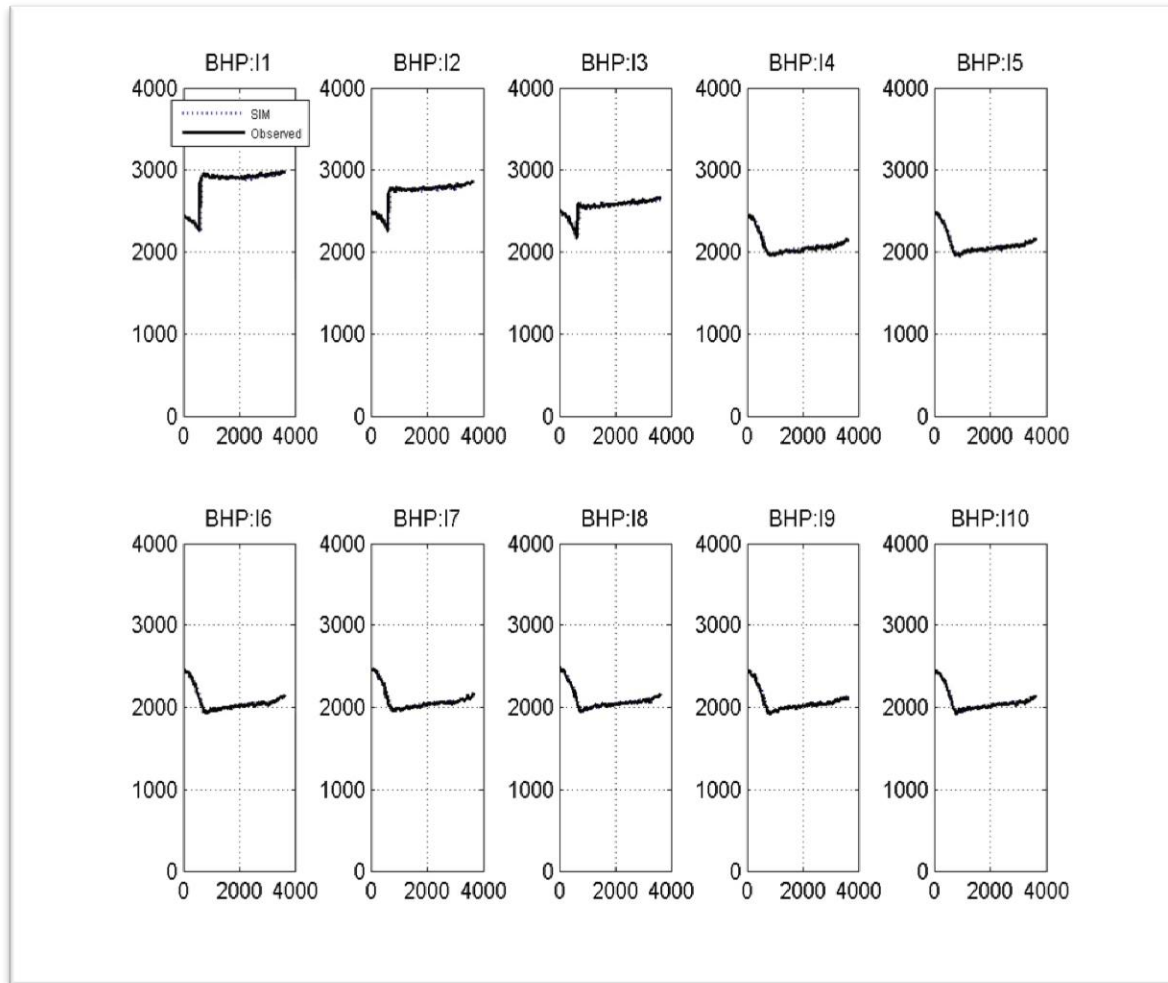
Fig.14—Bottomhole flowing pressure of water injector vs. time of observed data and simulation at end of MCMC chain.

44
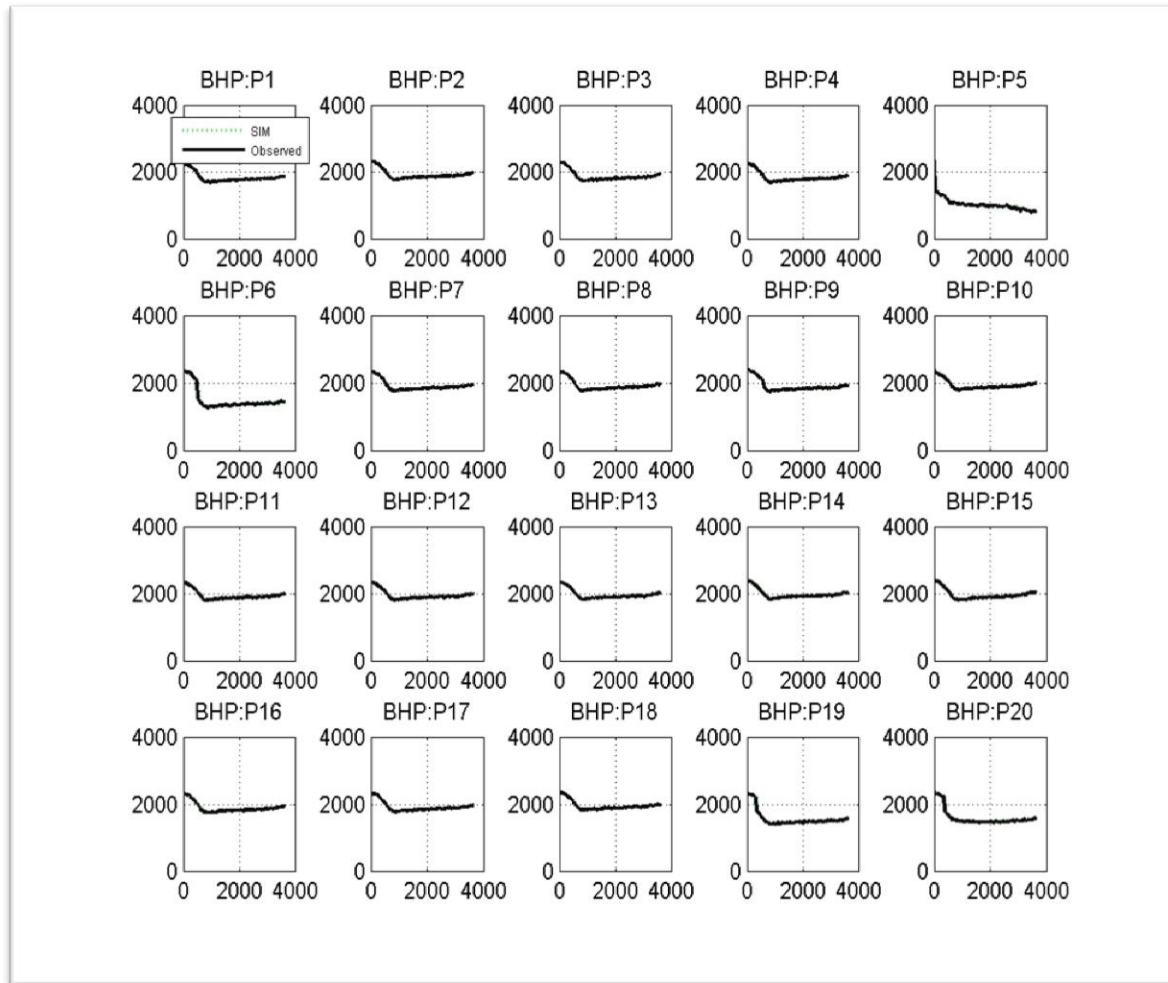
Fig. 15—Bottomhole flowing pressure of producer vs. time of observed data and simulation with at end of MCMC chain.
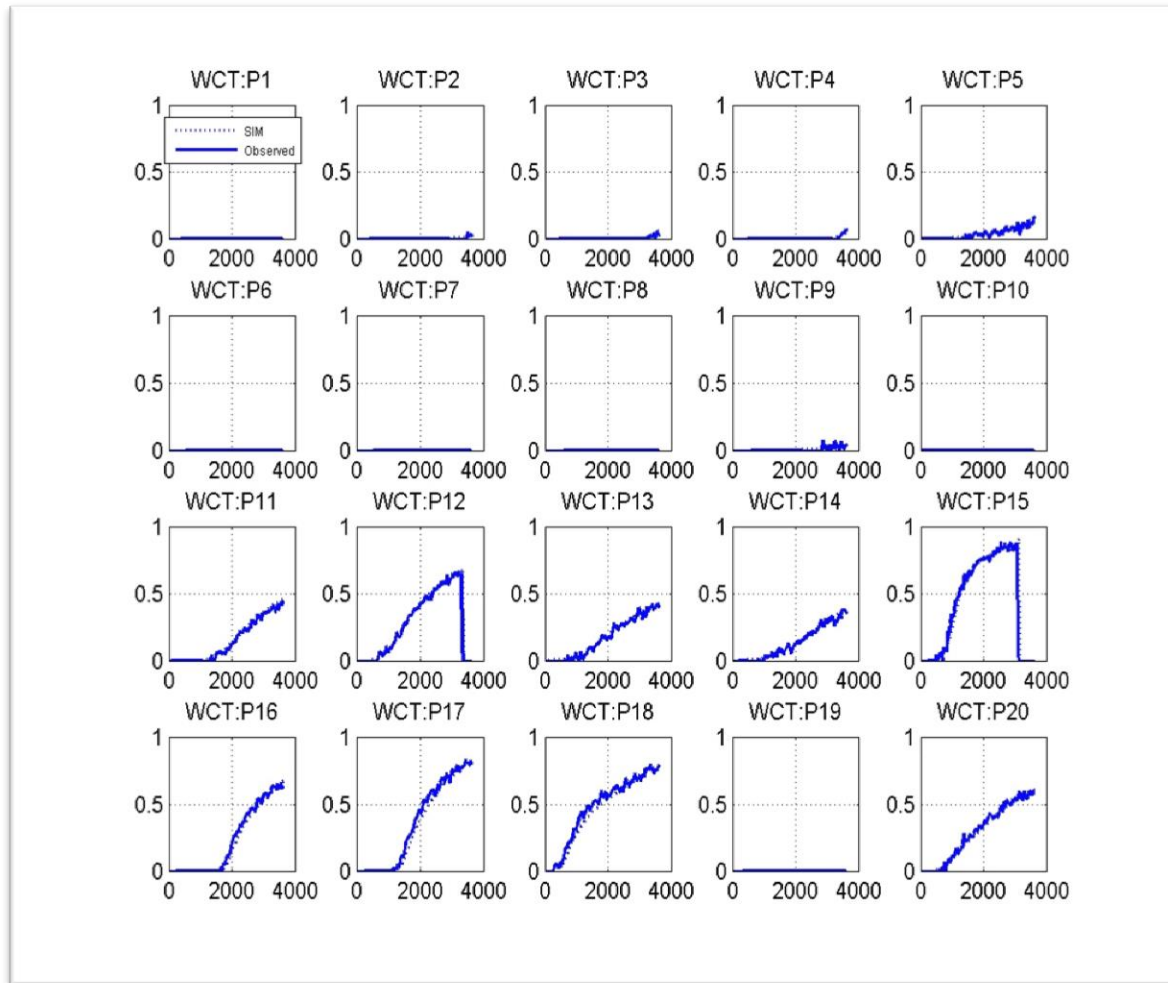
Fig. 16—Water cut at producer vs. time of observed data and simulation at end of MCMC

3.3.2 Distribution of Uncertain Parameters

The objective value of the chain seems to level off after about five hundred iterations; thus, the realizations accepted before that is excluded as a burn-in period (see Fig. 10). The total number of accepted realization is 773 models. The distribution of each uncertain parameter can be obtained from the accepted realizations after five hundred iterations. The cumulative distribution function of each uncertain parameter is shown in Fig. 17 and Fig. 18.
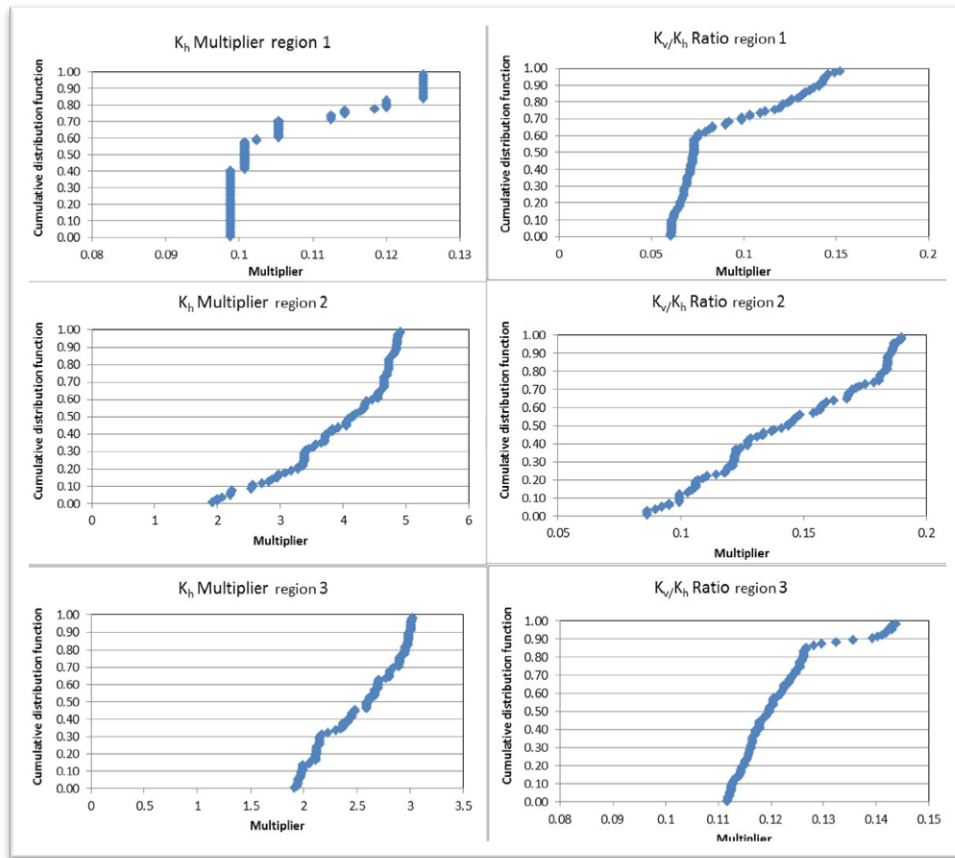
Fig. 17—Cumulative distribution function of k_h multiplier and k_v/k_h ratio for each zone.
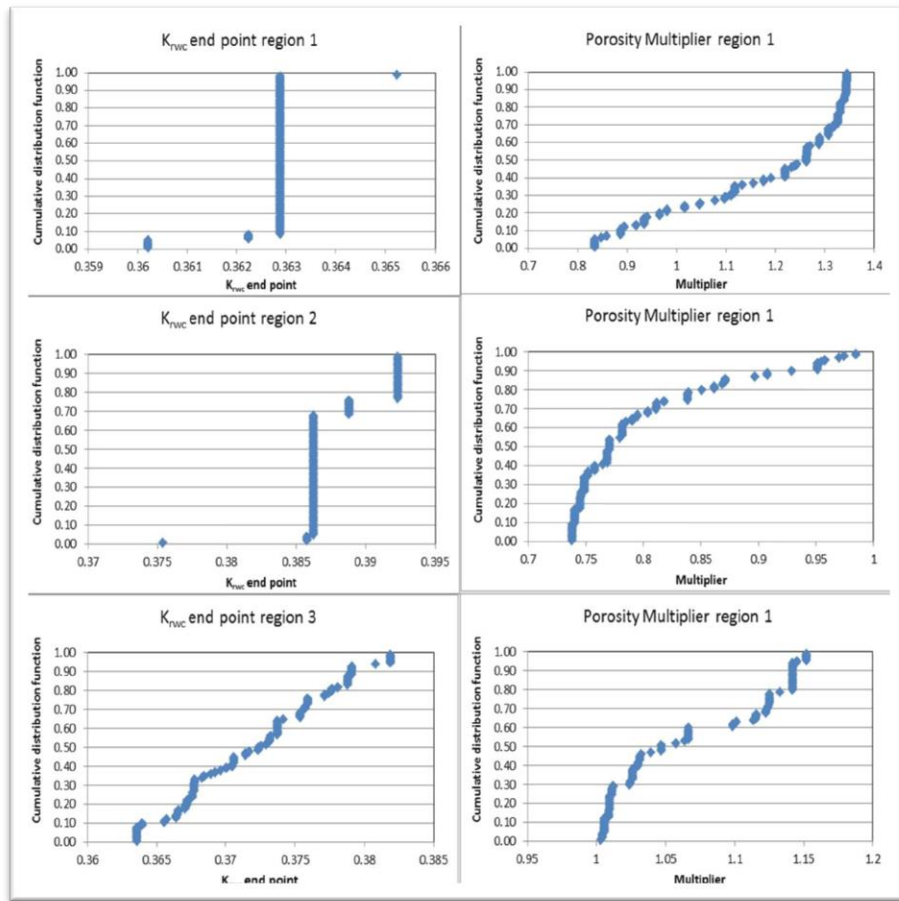
Fig. 18—Cumulative distribution function of k_{rwc} end point and ϕ multiplier for each zone.

The distributions of field original oil in place and oil in place for each region are also obtained (Fig. 19). We observe strong negative correlations between OOIP of Region 1 and OOIP of Regions 2 and 3 (Fig. 20). This is because in order to provide adequate reservoir energy to maintain bottomhole flowing pressure as seen in the observed data, the total reservoir volume needs to be maintained. Thus, if original oil in place in Region 1 decreases the original oil in place in Regions 2 and 3 increases. The distributions of all uncertain parameters are used in the model selection process.
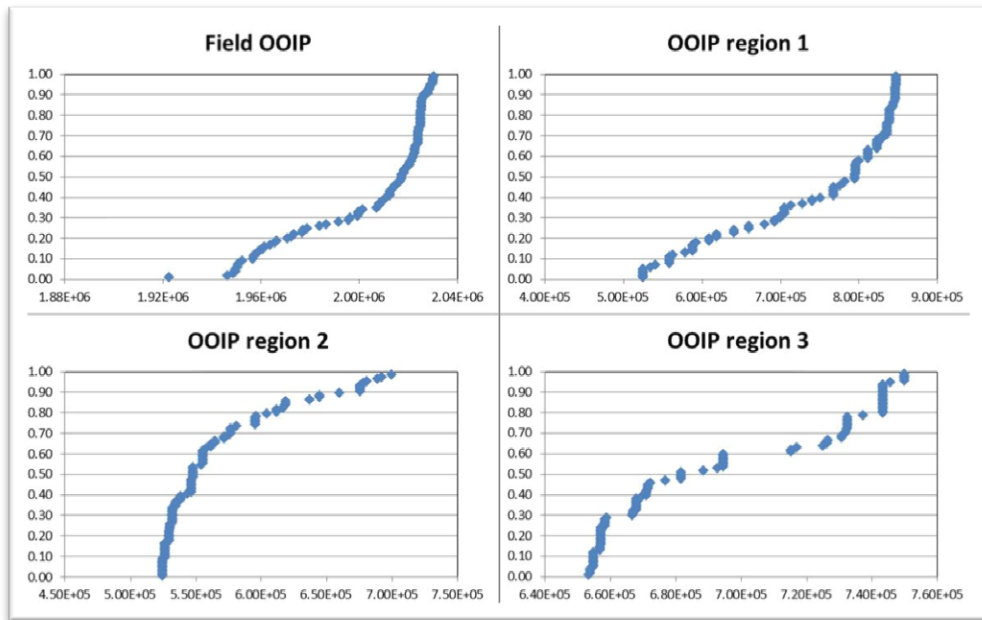
Fig. 19—Cumulative distribution functions of field OOIP and each region OOIP

49
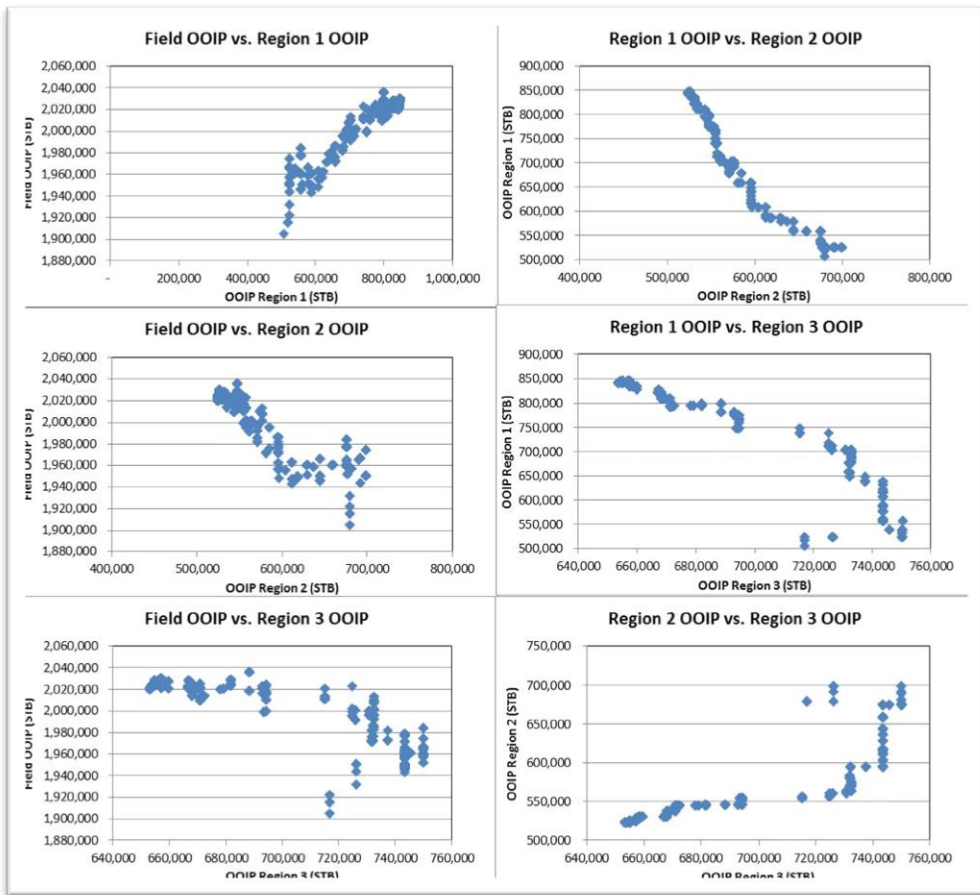
Fig. 20—Relationships between field OOIP and each of the region's OOIP.

4. MODEL SELECTION

4.1 Minimax Model Selection Method Implementation

Due to the large ensemble of models available (773), optimizing the development plan for all realizations is very computational expensive and impractical in this case. Thus, there is a need to select representative models to work with. The conventional model selection process relies on selecting target percentiles of some variables (Deutsch and Srinivasan (1996) and Odai and Ogbe (2011)). However, this can lead to sub-optimal representative models for optimizing the development plan because the selected models cannot represent the entire uncertainty space. Chen et al. (2013) proposed the Minimax selection method, which can select representative models that are close to target percentile and are maximally different (see Section 1.2.3.1). Due to difficulties in solving two optimization problems simultaneously, Chen et al. (2013) proposed a simplified version as a one-objective combinatorial optimization problem as follows:

$$\max_{X^1,\dots,X^P \in \blacksquare} \{ \min_{j,l;j \neq l} (\min_i |x_i^j - x_i^l|) \} \tag{17}$$

Subject to:

$$\sum_{j=l}^{P} \left\{ \max_k \min_l \left| \frac{y_k^j - \bar{y}_k^l}{\bar{y}_k^l} \right| \right\} \leq \varepsilon \, , \forall k = 1, \dots M; l = 1, \dots, P \tag{18}$$

$$unique\left([prctile(y_k^l), \dots, prctile(y_k^p)]\right) = P, \qquad \forall k = 1, \dots M \tag{19}$$

The simplified approach is suitable for small problems that have a few hundred models with fewer than ten input/output parameters and few target percentiles. The

51

simplified problem can be solved with exhaustive search. However, for larger problems the stochastic optimization algorithm can be utilized to speed up the calculation. In this study, the genetic algorithm is used in solving the simplified combinatorial optimization problem. The objective of Eqs. (17), (18) and (19) is to generate an ensemble of models that maximizes the minimum difference between any pair of input parameters and that matches the target percentiles within allowed tolerances. Steps to implement Minimax can be summarized as follows:

1. Categorize uncertain variables into two groups, input and output variables. Output variables are the variables that we would like to match the target percentiles. As model selection process is performed prior to optimization process, thus the net present value for each realization is not known. Field original oil in place and original regional oil in place are selected as output variables. Horizontal permeability multiplier, vertical permeability to horizontal permeability ratio, and end point of relative permeability to water are selected as input variables. Porosity multiplier is excluded as original oil in place is used as output parameter.

2. Select the number of representative models, target percentiles and tolerances. In this study, seven sets of ensemble are selected: (1) 3 models at percentiles 10th, 50th and 90th, (2) 3 models at percentiles 25th, 50th and 75th, (3) 5 models at percentiles 10th, 30th, 50th, 70th and 90th, (4) 9 models at every 10th percentile, (5) 19 models at every 5th percentile, (6) 33 models at every 3rd percentile, and (7) 49 models at every second percentile.

3. Apply the genetic algorithm to solve Eqs.(17), (18) and (19).

4.2 Model Selection Result

This section presents the results of using the Minimax model selection technique. As mentioned in Section 3.2, four sets of observed data are used in this study. The cumulative distribution function plots of input variables for the selected models in each ensemble are shown in Fig. 21 through Fig. 27 with the blue dot while other colors represent the selected realizations. The cumulative distribution function plots of output variables of the selected models in each ensemble are shown in Fig. 28 through Fig. 34 with the blue dot while other colors represent the selected realizations. By using the Minimax algorithm, we select the models that match the target percentiles and that are maximally different from each other.
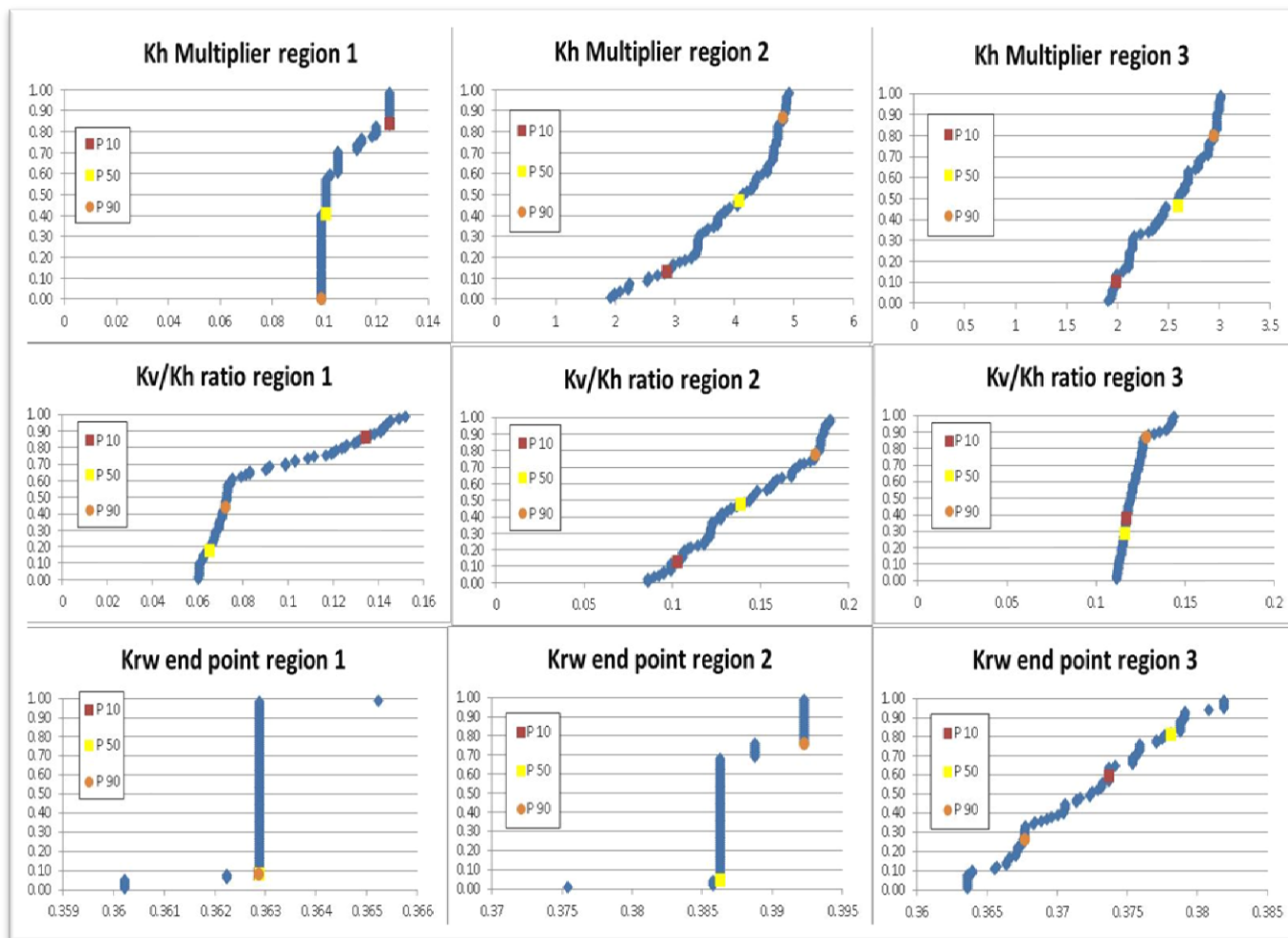
Fig. 21—Plot of selected models in Ensemble 1

54
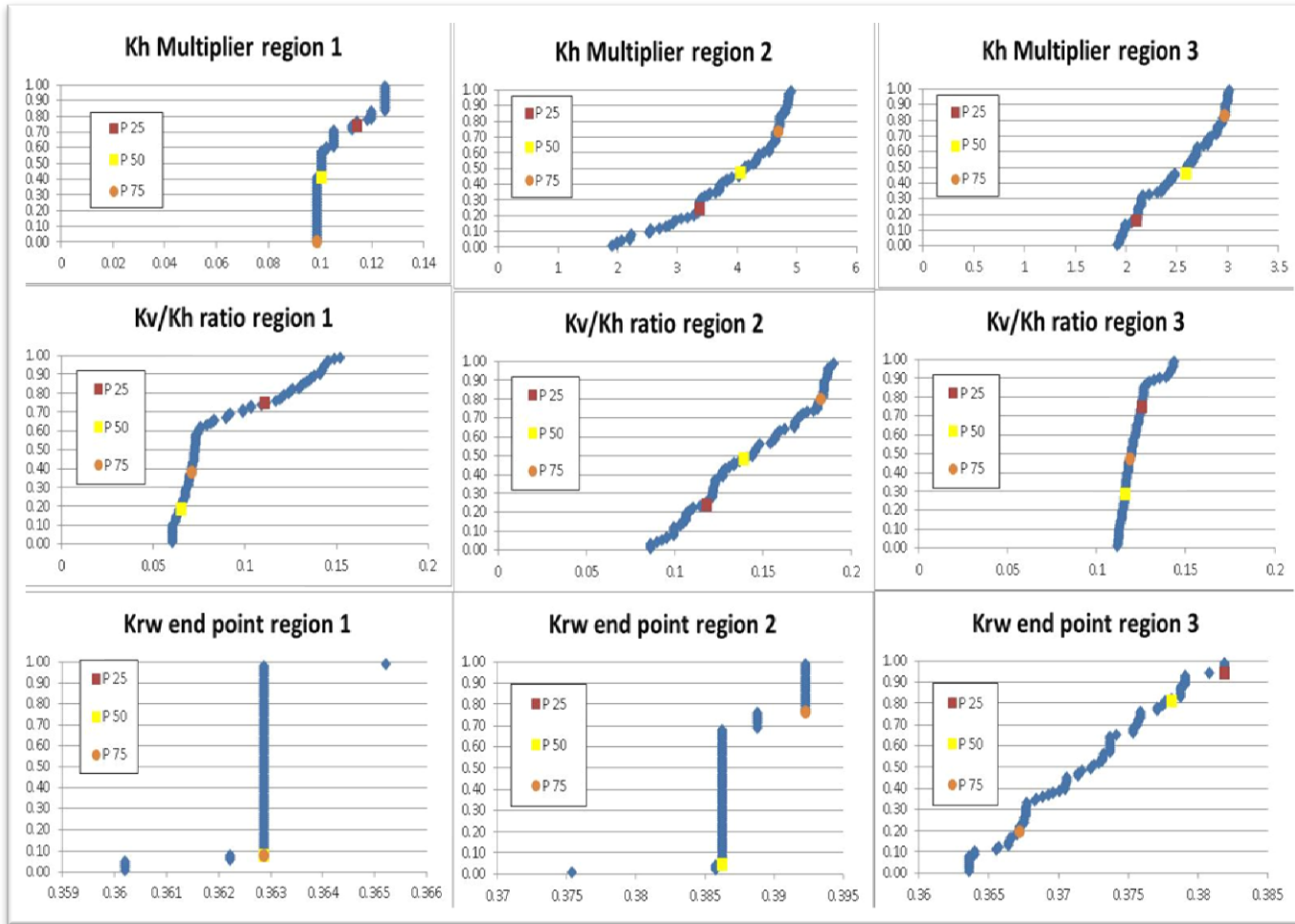
Fig. 22—Plot of selected models in Ensemble 2

55
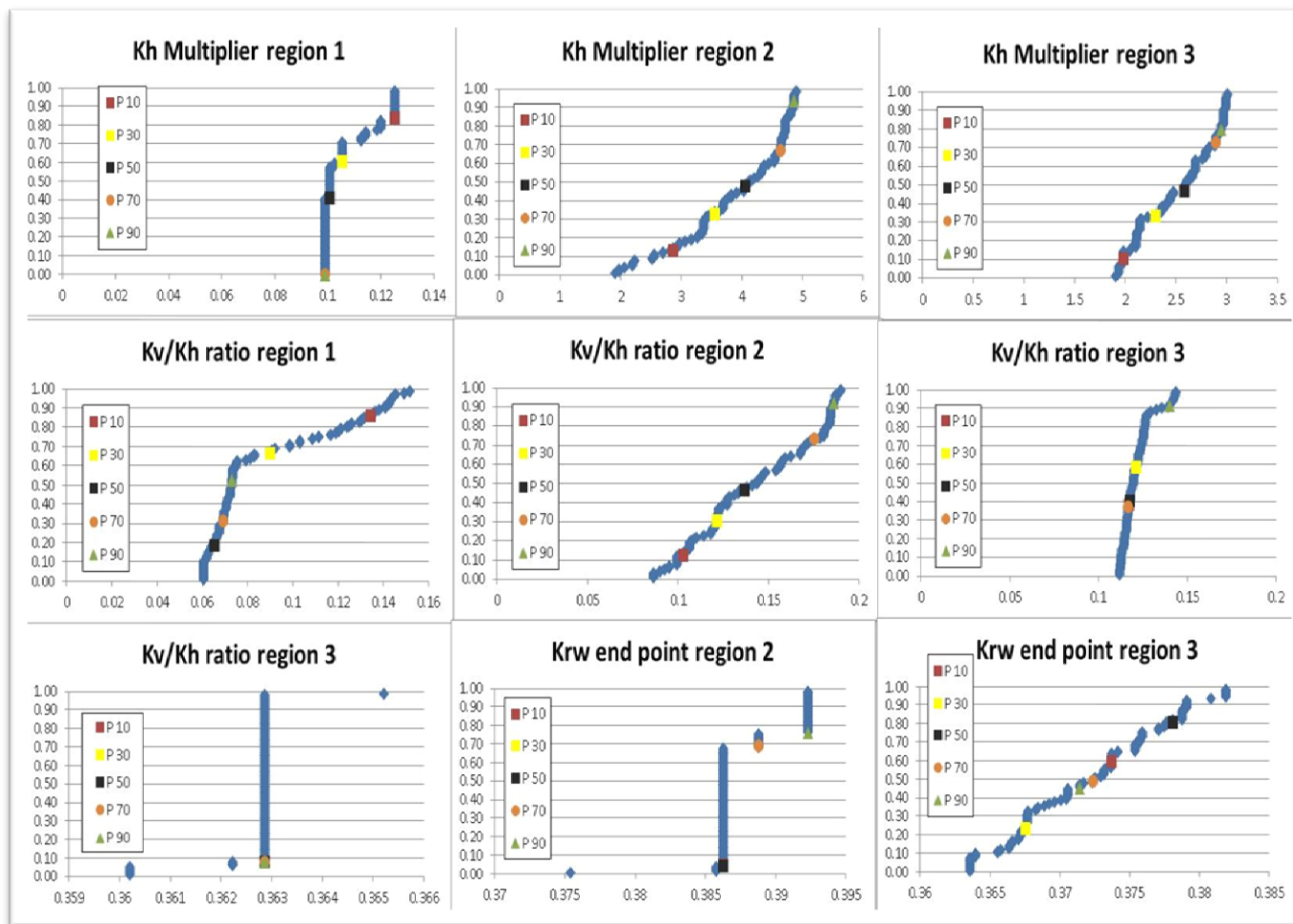
Fig. 23—Plot of selected models in Ensemble 3

56
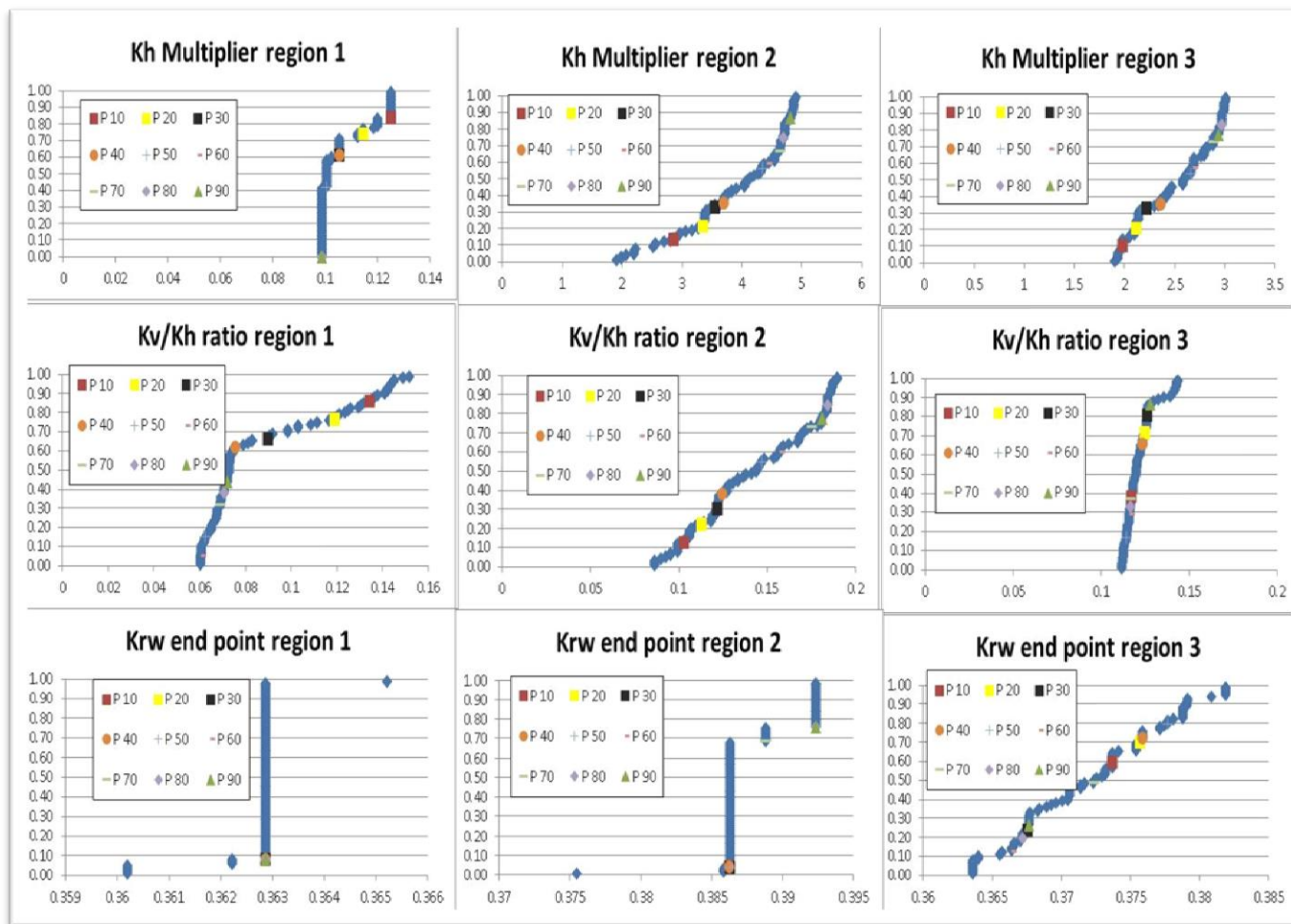
Fig. 24—Plot of selected models in Ensemble 4

57
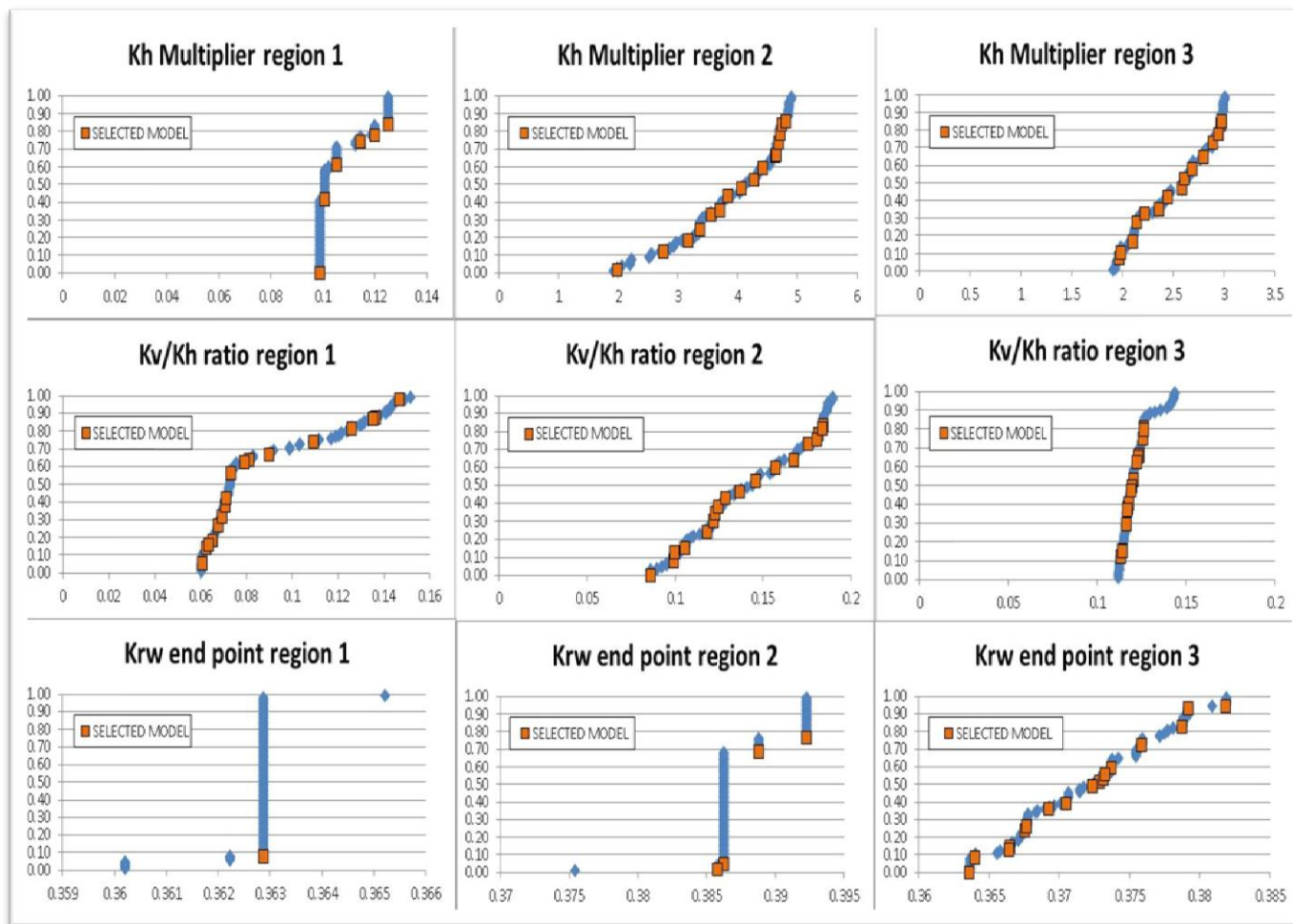
Fig. 25—Plot of selected models in Ensemble 5
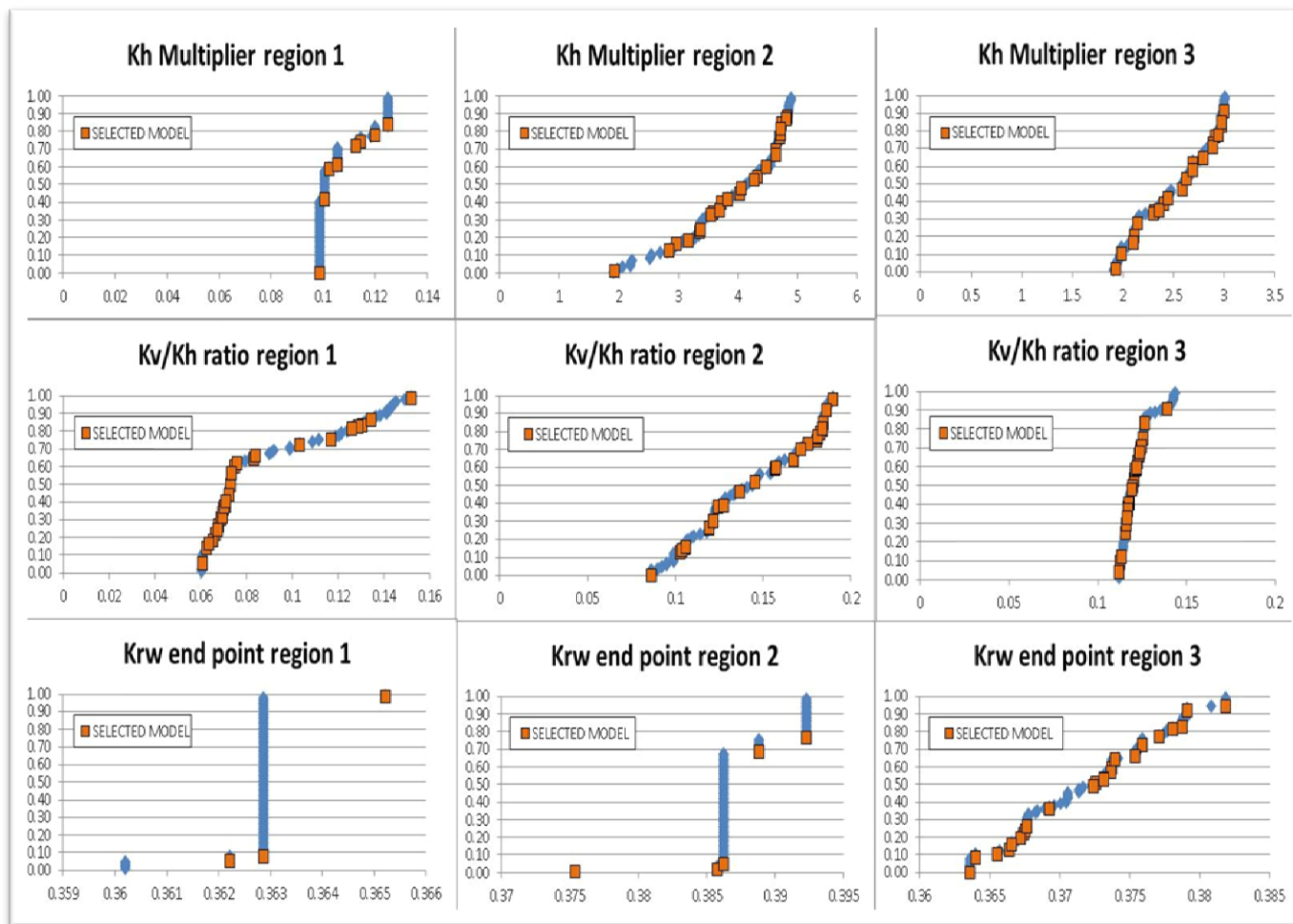
Fig. 26—Plot of selected models in Ensemble 6
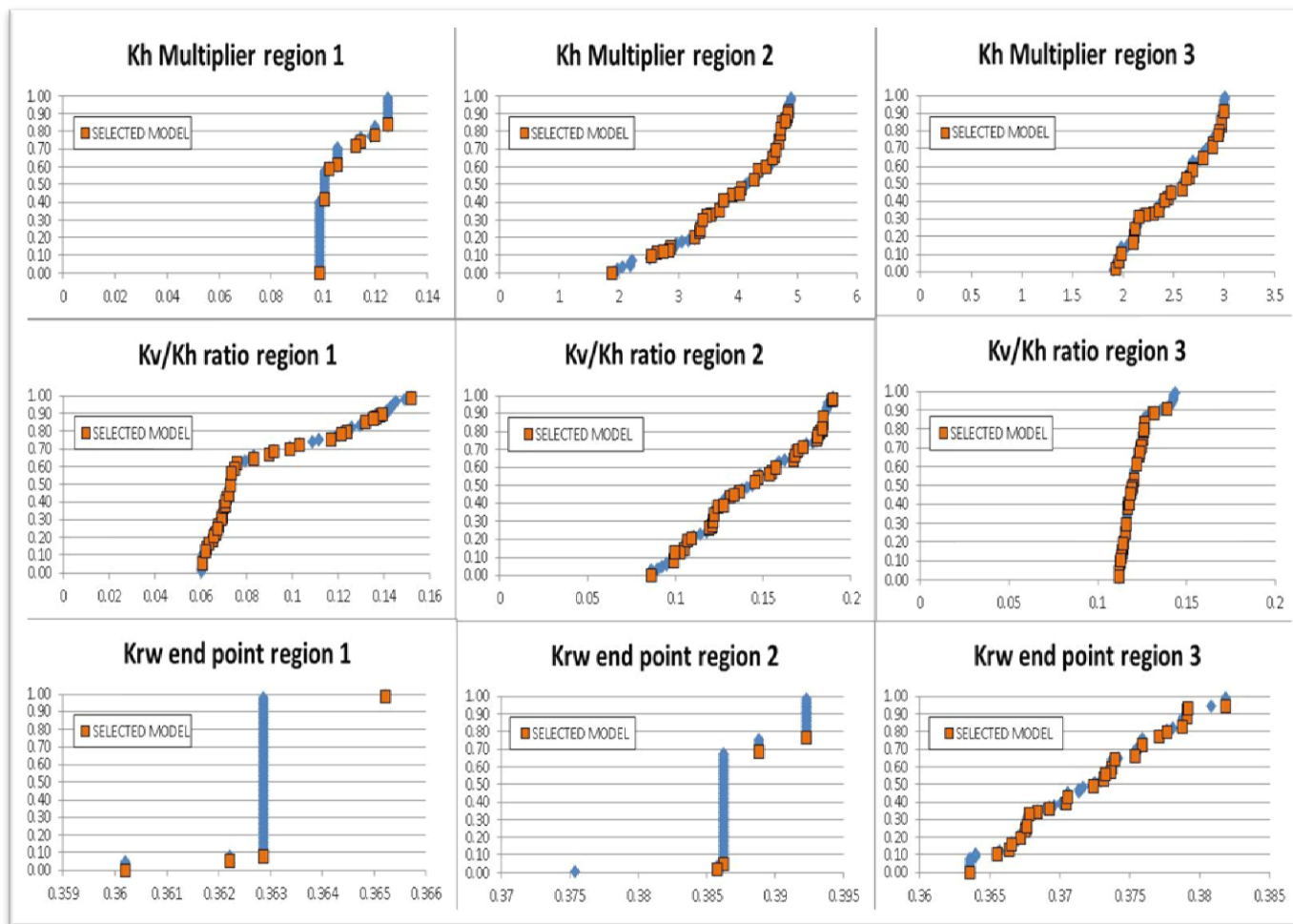
Fig. 27—Plot of selected models in Ensemble 7

60
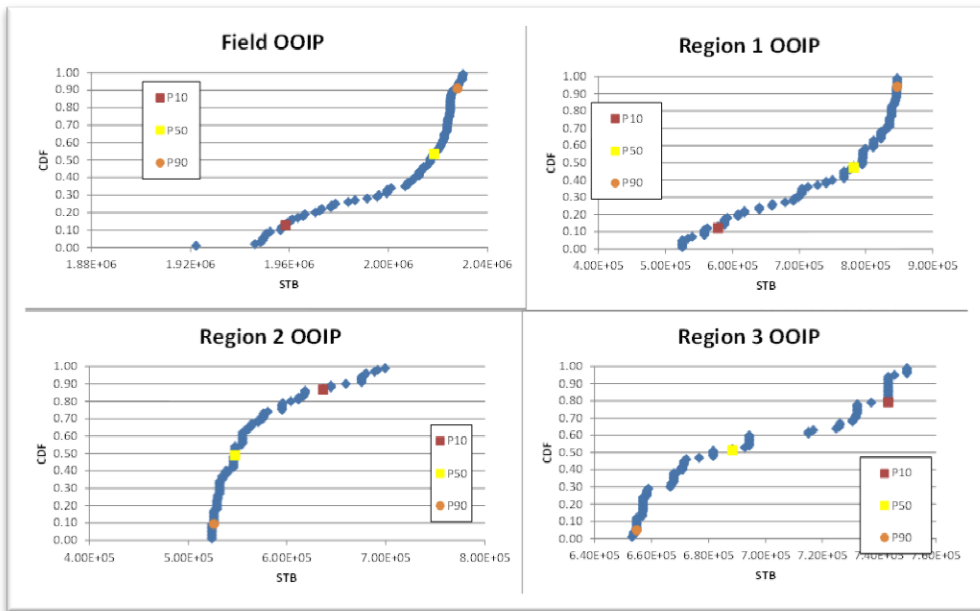
Fig. 28—Original oil in place of the selected models in Ensemble 1
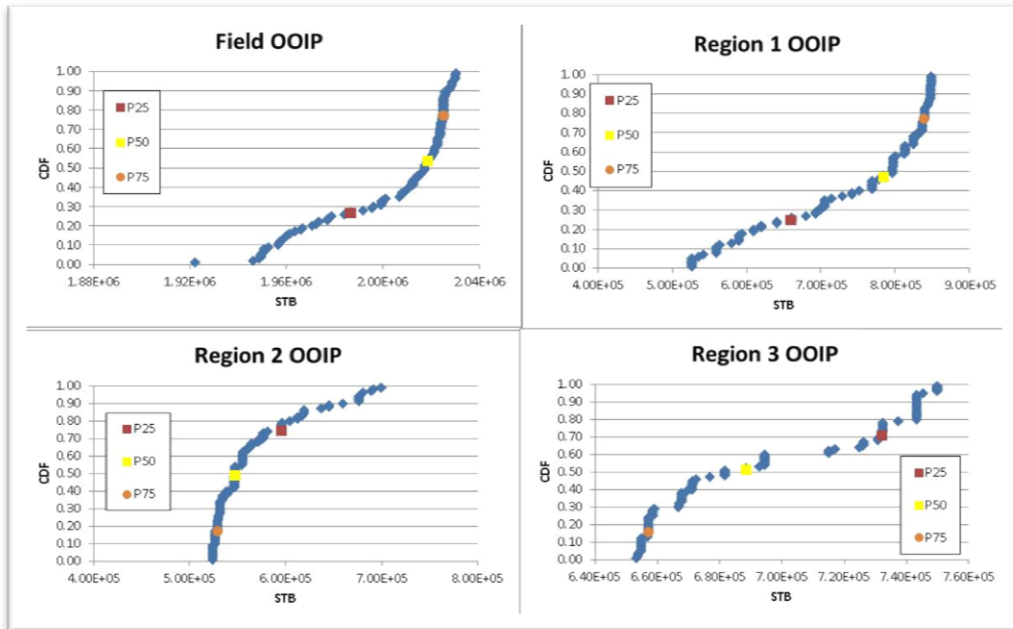
Fig. 29—Original oil in place of the selected models in Ensemble 2
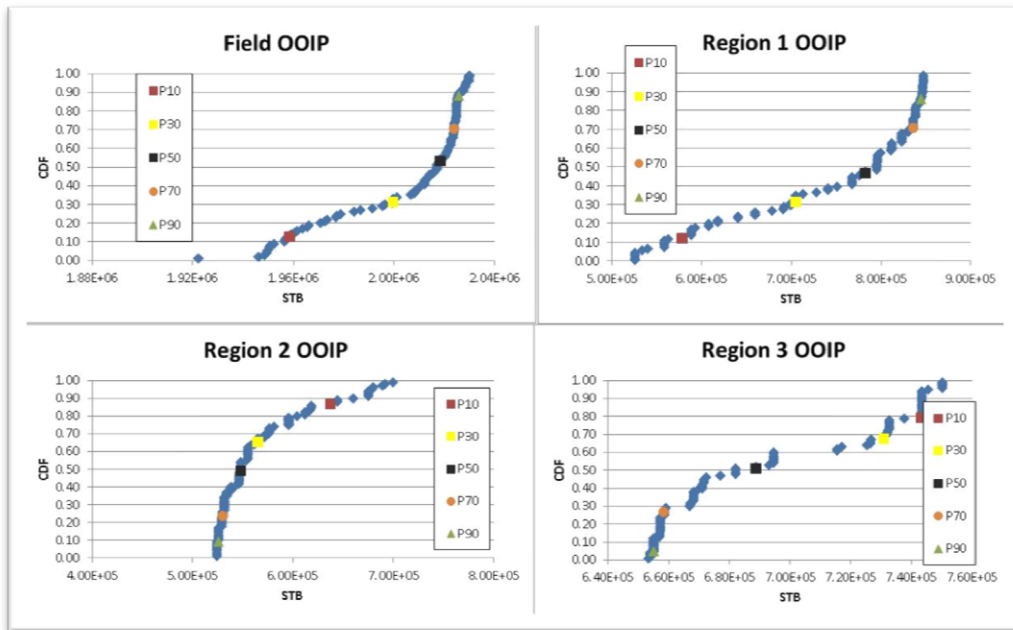
Fig. 30—Original oil in place of the selected models in Ensemble 3
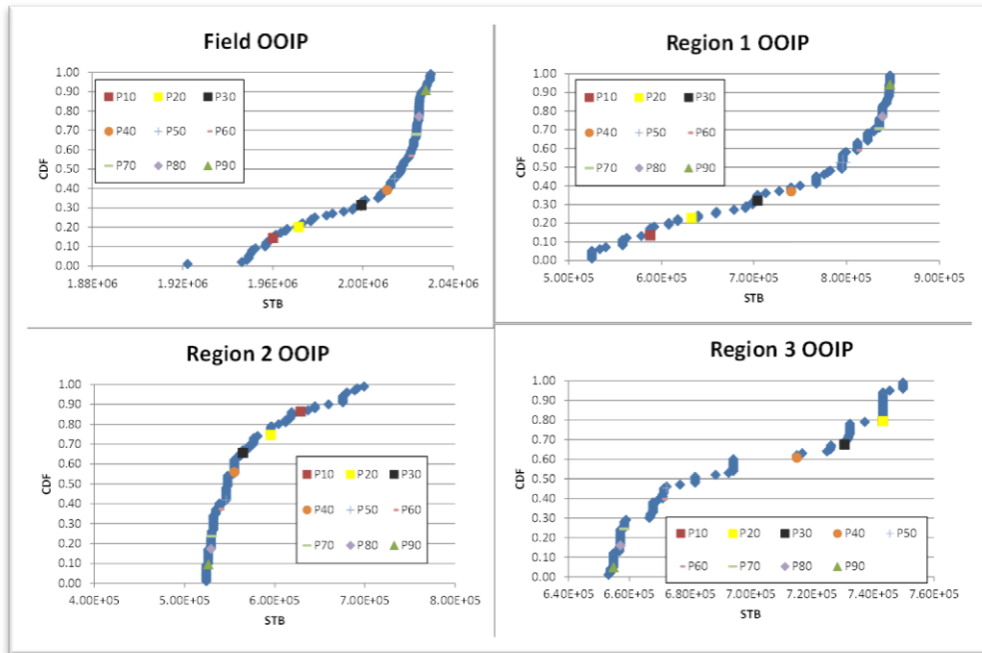
Fig. 31—Original oil in place of the selected models in Ensemble 4

62
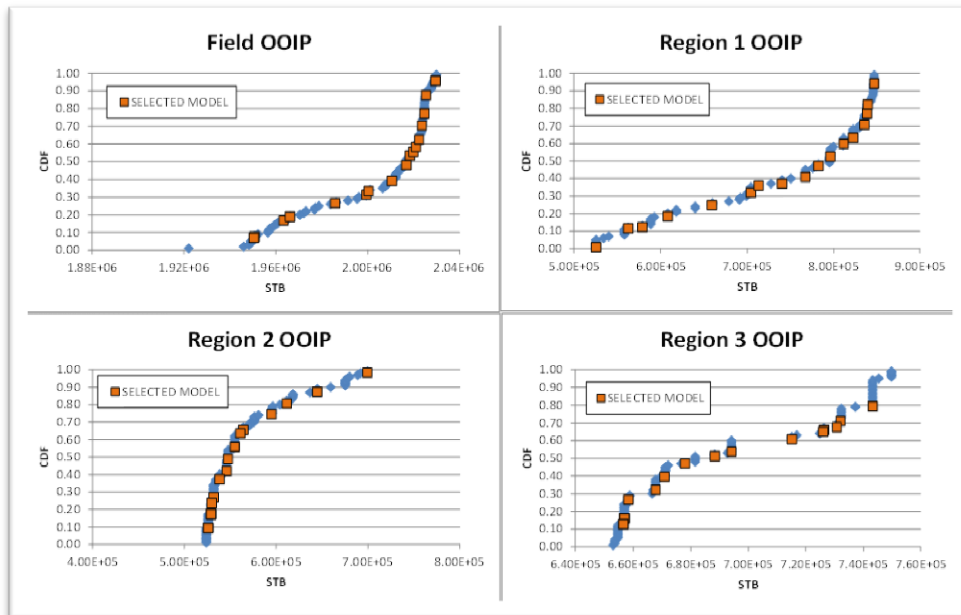
Fig. 32—Original oil in place of the selected models in Ensemble 5
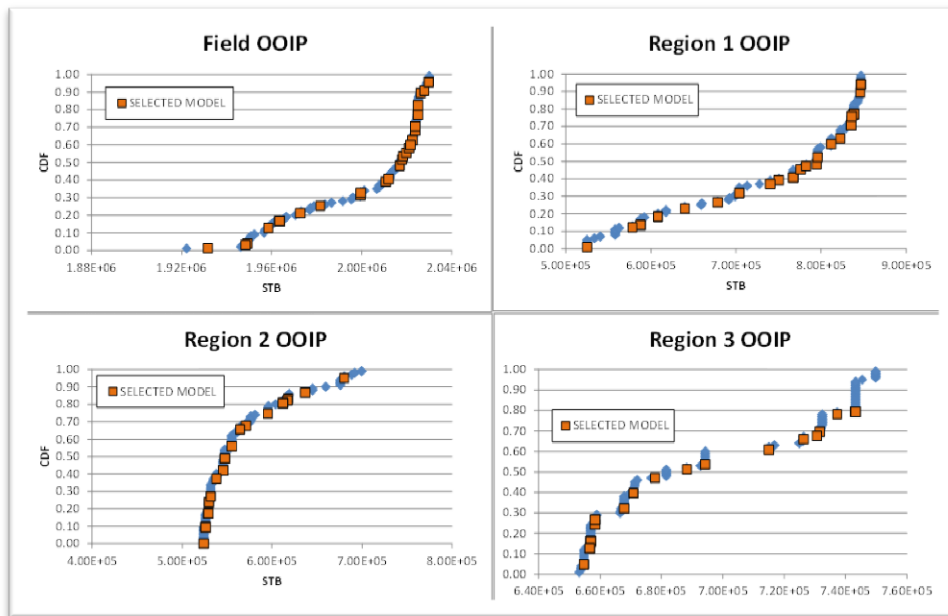
Fig. 33—Original oil in place of the selected models in Ensemble 6
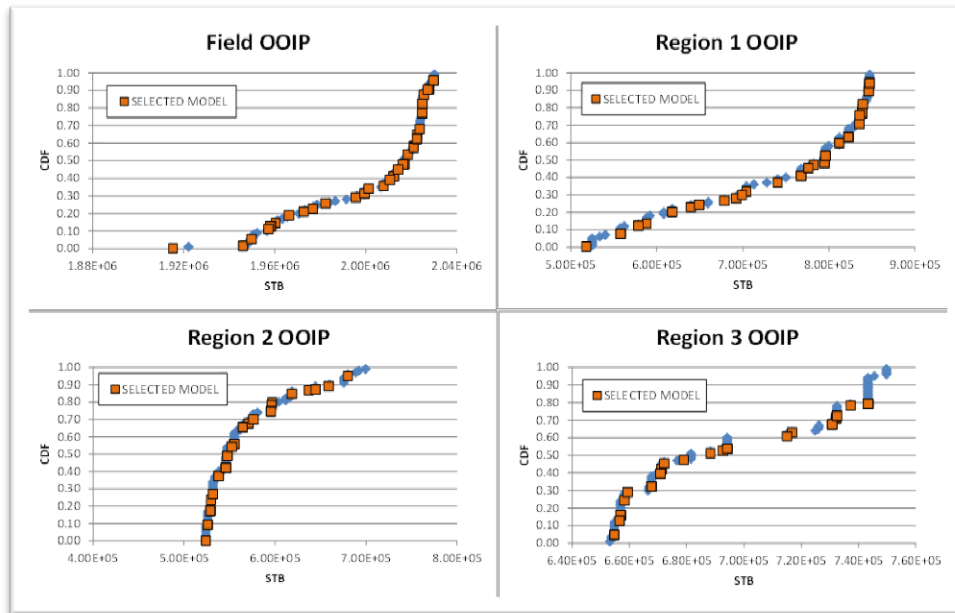
Fig. 34—Original oil in place of the selected models in Ensemble 7

4.3 Probability Weighting

Since each realization has a different probability of being true, each selected realization must be weighted differently during the optimization process. This study leverages the fact that we know the distribution of each uncertain parameter from the history matching process. Field original oil in place is chosen as the variable that will be used for calculating probabilities for selected realizations.

The field original oil in place distribution from the ensemble of 773 models is divided into twenty bins and the probability density of each bin is calculated (**Table 10**). Histogram of field original oil in place is shown in Fig. 35. For each smaller ensemble,

the probability weighting of each model in the ensemble is calculated via a table look up

from **Table 10** and then the probability of the entire ensemble is normalized to one.

Table 10—Field oil in place distribution

Lower FOIP Range(STB)	Upper FOIP Range(STB)	Frequency	Relative Frequency
-	1.905E+06	1	0.001
1.905E+06	1.912E+06	0	0.000
1.912E+06	1.919E+06	5	0.006
1.919E+06	1.926E+06	2	0.003
1.926E+06	1.933E+06	2	0.003
1.933E+06	1.940E+06	0	0.000
1.940E+06	1.946E+06	7	0.009
1.946E+06	1.953E+06	53	0.069
1.953E+06	1.960E+06	42	0.054
1.960E+06	1.967E+06	42	0.054
1.967E+06	1.974E+06	16	0.021
1.974E+06	1.981E+06	25	0.032
1.981E+06	1.988E+06	16	0.021
1.988E+06	1.995E+06	9	0.012
1.995E+06	2.002E+06	49	0.063
2.002E+06	2.009E+06	21	0.027
2.009E+06	2.016E+06	80	0.103
2.016E+06	2.023E+06	133	0.172
2.023E+06	2.030E+06	236	0.305
2.030E+06	2.037E+06	34	0.044
	Total	773	1
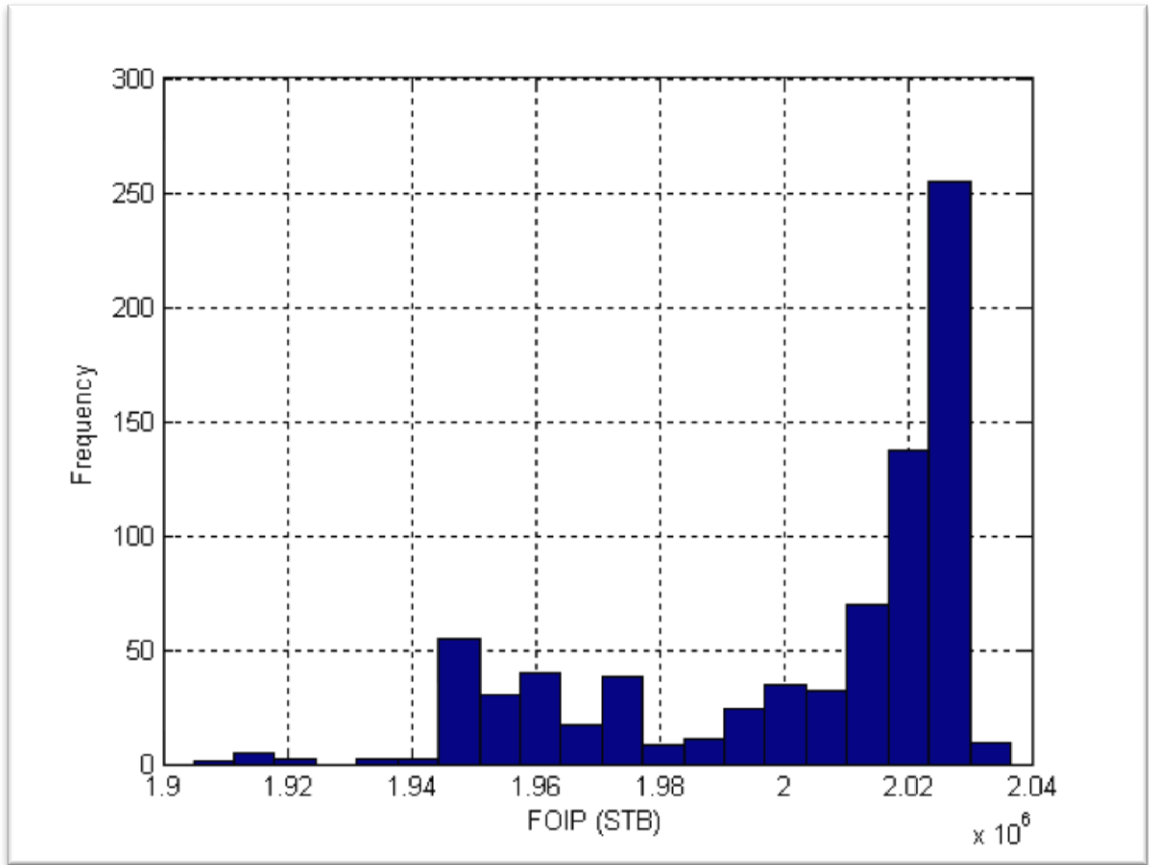
Fig. 35—Histogram of field oil in place

In this study, we will investigate optimizing different ensemble size including single most likely realization and will compare the resulting NPV from each scenario. The most-likely model used in this study is based on the maximum-a-posteriori (MAP) probability distribution. The descriptions of all selected realizations are summarized in **Table 11** and **Table 12**. **Table 11** contains the probability weighting, horizontal permeability multiplier and vertical to horizontal permeability ratio for each of the regions for all selected realizations in each ensemble. **Table 12** contains end point

relative permeability to water and porosity multiplier for each of the regions for all

selected realizations in each ensemble.

Table 11—Description of all selected realizations

Ensemble	Target Percentile	Realization	Field oil in place (STB)	Probability Weighting	k_h Multiplier region 1	k_h Multiplier region 2	k_h Multiplier region 3	k_v/k_h Ratio region 1	k_v/k_h Ratio region 2	k_v/k_h Ratio region 3
MAP	-	763	2.02E+06	-	0.099	4.894	3.010	0.073	0.187	0.143
1	10	103	1.96E+06	0.102	0.125	2.862	1.983	0.134	0.103	0.117
1	50	373	2.02E+06	0.324	0.101	4.063	2.584	0.065	0.139	0.116
1	90	683	2.03E+06	0.574	0.099	4.811	2.949	0.072	0.181	0.128
2	25	195	1.99E+06	0.042	0.114	3.371	2.102	0.111	0.118	0.125
2	50	373	2.02E+06	0.345	0.101	4.063	2.584	0.065	0.139	0.116
2	75	612	2.02E+06	0.613	0.099	4.692	2.971	0.071	0.183	0.119
3	10	103	1.96E+06	0.060	0.125	2.862	1.983	0.134	0.103	0.117
3	30	259	2.00E+06	0.070	0.105	3.550	2.298	0.090	0.122	0.121
3	50	371	2.02E+06	0.191	0.101	4.063	2.584	0.065	0.137	0.118
3	70	567	2.02E+06	0.339	0.099	4.642	2.890	0.069	0.175	0.117
3	90	701	2.03E+06	0.339	0.099	4.858	2.949	0.073	0.186	0.140
4	10	106	1.96E+06	0.038	0.125	2.862	1.983	0.134	0.103	0.117
4	20	176	1.97E+06	0.014	0.114	3.346	2.113	0.119	0.113	0.125
4	30	257	2.00E+06	0.044	0.105	3.550	2.218	0.090	0.122	0.126
4	40	294	2.01E+06	0.072	0.105	3.701	2.359	0.076	0.124	0.123
4	50	421	2.01E+06	0.072	0.101	4.342	2.658	0.062	0.147	0.114
4	60	462	2.02E+06	0.120	0.099	4.422	2.688	0.061	0.157	0.116
4	70	568	2.02E+06	0.213	0.099	4.642	2.890	0.069	0.175	0.117
4	80	605	2.02E+06	0.213	0.099	4.694	2.971	0.071	0.184	0.117
4	90	682	2.03E+06	0.213	0.099	4.811	2.942	0.072	0.181	0.128

Table 11—(Continued)

Ensemble	Target Percentile	Realization	Field oil in place (STB)	Probability Weighting	k_h Multiplier region 1	k_h Multiplier region 2	k_h Multiplier region 3	k_v/k_h Ratio region 1	k_v/k_h Ratio region 2	k_v/k_h Ratio region 3
5	5	17	1.95E+06	0.026	0.125	1.980	1.964	0.147	0.086	0.114
5	10	94	1.95E+06	0.026	0.125	2.748	1.983	0.136	0.099	0.117
5	15	144	1.96E+06	0.020	0.120	3.162	2.142	0.126	0.106	0.120
5	20	98	1.97E+06	0.020	0.125	2.748	1.983	0.135	0.100	0.117
5	25	197	1.99E+06	0.008	0.114	3.371	2.102	0.109	0.118	0.125
5	30	257	2.00E+06	0.024	0.105	3.550	2.218	0.090	0.122	0.126
5	35	278	2.00E+06	0.024	0.105	3.701	2.359	0.081	0.123	0.120
5	40	293	2.01E+06	0.039	0.105	3.701	2.359	0.079	0.124	0.123
5	45	337	2.02E+06	0.065	0.101	3.830	2.444	0.068	0.129	0.118
5	50	369	2.02E+06	0.065	0.101	4.063	2.584	0.065	0.137	0.118
5	55	410	2.02E+06	0.065	0.101	4.269	2.605	0.062	0.146	0.113
5	60	463	2.02E+06	0.065	0.099	4.422	2.688	0.061	0.157	0.116
5	65	517	2.02E+06	0.065	0.099	4.631	2.794	0.063	0.167	0.114
5	70	567	2.02E+06	0.115	0.099	4.642	2.890	0.069	0.175	0.117
5	75	614	2.02E+06	0.115	0.099	4.692	2.971	0.071	0.182	0.119
5	80	615	2.02E+06	0.115	0.099	4.692	2.971	0.071	0.182	0.123
5	85	637	2.03E+06	0.115	0.099	4.718	2.972	0.071	0.181	0.126
5	90	660	2.03E+06	0.016	0.099	4.735	2.972	0.073	0.184	0.123
5	95	668	2.03E+06	0.016	0.099	4.806	2.942	0.073	0.184	0.122
6	2	9	1.93E+06	0.001	0.125	1.920	1.930	0.152	0.086	0.113

Table 11—(Continued)

Ensemble	Target Percentile	Realization	Field oil in place (STB)	Probability Weighting	k_h Multiplier region 1	k_h Multiplier region 2	k_h Multiplier region 3	k_v/k_h Ratio region 1	k_v/k_h Ratio region 2	k_v/k_h Ratio region 3
6	5	130	1.95E+06	0.014	0.120	2.965	2.142	0.130	0.106	0.120
6	8	132	1.95E+06	0.014	0.120	2.965	2.142	0.129	0.106	0.120
6	11	102	1.96E+06	0.011	0.125	2.849	1.983	0.134	0.103	0.117
6	14	144	1.96E+06	0.011	0.120	3.162	2.142	0.126	0.106	0.120
6	17	145	1.96E+06	0.011	0.120	3.162	2.142	0.126	0.103	0.120
6	20	143	1.96E+06	0.011	0.120	3.162	2.142	0.126	0.106	0.120
6	23	186	1.97E+06	0.004	0.114	3.363	2.113	0.117	0.119	0.125
6	26	210	1.98E+06	0.004	0.112	3.371	2.102	0.103	0.119	0.125
6	29	266	2.00E+06	0.013	0.105	3.588	2.312	0.083	0.122	0.120
6	32	261	2.00E+06	0.013	0.105	3.550	2.298	0.084	0.122	0.121
6	35	310	2.00E+06	0.013	0.102	3.727	2.406	0.075	0.127	0.122
6	38	294	2.01E+06	0.021	0.105	3.701	2.359	0.076	0.124	0.123
6	41	448	2.01E+06	0.021	0.101	4.325	2.692	0.061	0.157	0.116
6	44	329	2.02E+06	0.034	0.101	3.823	2.444	0.068	0.127	0.122
6	47	324	2.02E+06	0.034	0.101	3.823	2.444	0.070	0.127	0.124
6	50	361	2.02E+06	0.034	0.101	4.035	2.584	0.066	0.137	0.118
6	53	371	2.02E+06	0.034	0.101	4.063	2.584	0.065	0.137	0.118
6	56	408	2.02E+06	0.034	0.101	4.269	2.626	0.062	0.145	0.113
6	59	520	2.02E+06	0.034	0.099	4.645	2.794	0.063	0.167	0.112
6	62	465	2.02E+06	0.034	0.099	4.480	2.688	0.061	0.157	0.116

Table 11—(Continued)

Ensemble	Target Percentile	Realization	Field oil in place (STB)	Probability Weighting	k_h Multiplier region 1	k_h Multiplier region 2	k_h Multiplier region 3	k_v/k_h Ratio region 1	k_v/k_h Ratio region 2	k_v/k_h Ratio region 3
6	65	595	2.02E+06	0.034	0.099	4.712	2.907	0.069	0.184	0.122
6	68	580	2.02E+06	0.061	0.099	4.736	2.890	0.069	0.181	0.119
6	71	567	2.02E+06	0.061	0.099	4.642	2.890	0.069	0.175	0.117
6	74	621	2.02E+06	0.061	0.099	4.718	2.972	0.071	0.184	0.124
6	77	618	2.02E+06	0.061	0.099	4.718	2.971	0.071	0.184	0.124
6	80	627	2.02E+06	0.061	0.099	4.718	2.972	0.070	0.181	0.126
6	83	556	2.03E+06	0.061	0.099	4.642	2.884	0.067	0.171	0.117
6	86	730	2.03E+06	0.061	0.099	4.827	2.994	0.073	0.190	0.139
6	89	678	2.03E+06	0.061	0.099	4.811	2.942	0.072	0.182	0.123
6	92	679	2.03E+06	0.061	0.099	4.811	2.942	0.072	0.181	0.123
6	95	675	2.03E+06	0.009	0.099	4.811	2.942	0.073	0.184	0.123
6	98	653	2.03E+06	0.009	0.099	4.720	2.972	0.071	0.186	0.123
7	2	85	1.95E+06	0.001	0.125	2.551	1.961	0.138	0.099	0.114
7	4	88	1.95E+06	0.001	0.125	2.642	1.961	0.136	0.099	0.114
7	6	6	1.92E+06	0.001	0.125	1.884	1.930	0.152	0.086	0.113
7	8	113	1.95E+06	0.010	0.125	2.869	1.983	0.132	0.105	0.119
7	10	82	1.96E+06	0.008	0.125	2.551	1.961	0.139	0.099	0.115
7	12	101	1.96E+06	0.008	0.125	2.849	1.983	0.135	0.103	0.117
7	14	157	1.96E+06	0.008	0.120	3.276	2.113	0.124	0.107	0.120
7	16	161	1.96E+06	0.008	0.120	3.276	2.113	0.121	0.109	0.122
7	18	191	1.98E+06	0.001	0.114	3.371	2.102	0.117	0.120	0.125
7	20	98	1.97E+06	0.008	0.125	2.748	1.983	0.135	0.100	0.117

Table 11—(Continued)

Ensemble	Target Percentile	Realization	Field oil in place (STB)	Probability Weighting	k_h Multiplier region 1	k_h Multiplier region 2	k_h Multiplier region 3	k_v/k_h Ratio region 1	k_v/k_h Ratio region 2	k_v/k_h Ratio region 3
7	22	187	1.97E+06	0.003	0.114	3.363	2.102	0.117	0.119	0.125
7	24	190	1.98E+06	0.001	0.114	3.371	2.102	0.117	0.119	0.125
7	26	211	1.98E+06	0.003	0.112	3.371	2.102	0.103	0.119	0.125
7	28	263	2.00E+06	0.009	0.105	3.550	2.298	0.083	0.122	0.120
7	30	222	2.00E+06	0.009	0.105	3.371	2.120	0.099	0.121	0.131
7	32	264	2.00E+06	0.009	0.105	3.550	2.298	0.083	0.122	0.120
7	34	255	2.00E+06	0.009	0.105	3.480	2.218	0.090	0.122	0.126
7	36	241	2.01E+06	0.004	0.105	3.404	2.160	0.092	0.122	0.126
7	38	444	2.01E+06	0.014	0.101	4.348	2.658	0.061	0.155	0.116
7	40	441	2.01E+06	0.014	0.101	4.348	2.658	0.061	0.154	0.116
7	42	294	2.01E+06	0.014	0.105	3.701	2.359	0.076	0.124	0.123
7	44	340	2.02E+06	0.024	0.101	3.911	2.444	0.068	0.131	0.118
7	46	318	2.02E+06	0.024	0.102	3.764	2.411	0.075	0.127	0.124
7	48	362	2.02E+06	0.024	0.101	4.035	2.584	0.066	0.137	0.118
7	50	370	2.02E+06	0.024	0.101	4.063	2.584	0.065	0.137	0.118
7	52	416	2.01E+06	0.014	0.101	4.269	2.654	0.062	0.147	0.114
7	54	516	2.02E+06	0.024	0.099	4.631	2.794	0.062	0.167	0.114
7	56	354	2.02E+06	0.024	0.101	4.035	2.474	0.067	0.133	0.118
7	58	467	2.02E+06	0.024	0.099	4.480	2.690	0.061	0.157	0.116
7	60	465	2.02E+06	0.024	0.099	4.480	2.688	0.061	0.157	0.116
7	62	514	2.02E+06	0.024	0.099	4.631	2.794	0.062	0.169	0.113
7	64	407	2.02E+06	0.024	0.101	4.269	2.626	0.062	0.145	0.113

Table 11—(Continued)

Ensemble	Target Percentile	Realization	Field oil in place (STB)	Probability Weighting	k_h Multiplier region 1	k_h Multiplier region 2	k_h Multiplier region 3	k_v/k_h Ratio region 1	k_v/k_h Ratio region 2	k_v/k_h Ratio region 3
7	66	520	2.02E+06	0.024	0.099	4.645	2.794	0.063	0.167	0.112
7	68	524	2.02E+06	0.024	0.099	4.645	2.794	0.065	0.168	0.113
7	70	575	2.02E+06	0.043	0.099	4.753	2.890	0.069	0.181	0.119
7	72	618	2.02E+06	0.043	0.099	4.718	2.971	0.071	0.184	0.124
7	74	511	2.02E+06	0.024	0.099	4.583	2.794	0.062	0.170	0.114
7	76	603	2.02E+06	0.043	0.099	4.694	2.971	0.069	0.184	0.118
7	78	625	2.02E+06	0.043	0.099	4.718	2.972	0.070	0.183	0.126
7	80	620	2.02E+06	0.043	0.099	4.718	2.971	0.071	0.184	0.124
7	82	615	2.02E+06	0.043	0.099	4.692	2.971	0.071	0.182	0.123
7	84	637	2.03E+06	0.043	0.099	4.718	2.972	0.071	0.181	0.126
7	86	726	2.03E+06	0.043	0.099	4.841	2.994	0.073	0.189	0.139
7	88	676	2.03E+06	0.043	0.099	4.811	2.942	0.073	0.184	0.123
7	90	654	2.03E+06	0.006	0.099	4.720	2.972	0.071	0.184	0.123
7	92	686	2.03E+06	0.043	0.099	4.848	2.949	0.072	0.181	0.132
7	94	733	2.03E+06	0.043	0.099	4.842	2.994	0.073	0.190	0.139
7	96	549	2.03E+06	0.043	0.099	4.645	2.884	0.067	0.173	0.112
7	98	667	2.03E+06	0.006	0.099	4.806	2.942	0.073	0.184	0.122

Table 12—Additional description of all selected realizations

Ensemble	Target Percentile	Realization	k_{rw} end point region 1	k_{rw} end point region 2	k_{rw} end point region 3	Porosity Multiplier region 1	Porosity Multiplier region 2	Porosity Multiplier region 3
MAP	-	763	0.360	0.392	0.372	1.341	0.738	1.004
1	10	103	0.363	0.386	0.374	0.919	0.897	1.142
1	50	373	0.363	0.386	0.378	1.243	0.771	1.057
1	90	683	0.363	0.392	0.368	1.346	0.741	1.006
2	25	195	0.363	0.386	0.382	1.047	0.838	1.124
2	50	373	0.363	0.386	0.378	1.243	0.771	1.057
2	75	612	0.363	0.392	0.367	1.332	0.745	1.009
3	10	103	0.363	0.386	0.374	0.919	0.897	1.142
3	30	259	0.363	0.386	0.368	1.119	0.795	1.122
3	50	371	0.363	0.386	0.378	1.243	0.771	1.057
3	70	567	0.363	0.389	0.372	1.327	0.746	1.011
3	90	701	0.363	0.392	0.371	1.342	0.741	1.006
4	10	106	0.363	0.386	0.374	0.934	0.885	1.142
4	20	176	0.363	0.386	0.376	1.005	0.838	1.142
4	30	257	0.363	0.386	0.368	1.119	0.795	1.122
4	40	294	0.363	0.386	0.376	1.176	0.781	1.098
4	50	421	0.363	0.386	0.378	1.265	0.769	1.032
4	60	462	0.363	0.386	0.366	1.289	0.758	1.030
4	70	568	0.363	0.389	0.372	1.327	0.746	1.011
4	80	605	0.363	0.392	0.367	1.332	0.745	1.009
4	90	682	0.363	0.392	0.368	1.346	0.741	1.006

Table 12—(Continued)

Ensemble	Target Percentile	Realization	k_rw end point region 1	k_rw end point region 2	k_rw end point region 3	Porosity Multiplier region 1	Porosity Multiplier region 2	Porosity Multiplier region 3
5	5	17	0.363	0.386	0.364	0.834	0.984	1.115
5	10	94	0.363	0.386	0.370	0.894	0.908	1.142
5	15	144	0.363	0.386	0.364	0.966	0.862	1.142
5	20	98	0.363	0.386	0.374	0.919	0.908	1.142
5	25	197	0.363	0.386	0.382	1.047	0.837	1.124
5	30	257	0.363	0.386	0.368	1.119	0.795	1.122
5	35	278	0.363	0.386	0.373	1.133	0.790	1.115
5	40	293	0.363	0.386	0.376	1.176	0.781	1.098
5	45	337	0.363	0.386	0.366	1.219	0.781	1.066
5	50	369	0.363	0.386	0.379	1.243	0.771	1.057
5	55	410	0.363	0.386	0.379	1.265	0.769	1.041
5	60	463	0.363	0.386	0.366	1.290	0.758	1.030
5	65	517	0.363	0.386	0.373	1.307	0.749	1.026
5	70	567	0.363	0.389	0.372	1.327	0.746	1.011
5	75	614	0.363	0.392	0.369	1.332	0.745	1.009
5	80	615	0.363	0.392	0.369	1.332	0.745	1.009
5	85	637	0.363	0.392	0.373	1.334	0.745	1.009
5	90	660	0.363	0.392	0.368	1.346	0.741	1.009
5	95	668	0.363	0.392	0.368	1.346	0.741	1.009

Table 12—(Continued)

Ensemble	Target Percentile	Realization	k_{rw} end point region 1	k_{rw} end point region 2	k_{rw} end point region 3	Porosity Multiplier region 1	Porosity Multiplier region 2	Porosity Multiplier region 3
6	2	9	0.365	0.386	0.364	0.834	0.958	1.115
6	5	130	0.363	0.386	0.367	0.934	0.870	1.142
6	8	132	0.363	0.386	0.367	0.934	0.869	1.142
6	11	102	0.363	0.386	0.374	0.919	0.897	1.142
6	14	144	0.363	0.386	0.364	0.966	0.862	1.142
6	17	145	0.363	0.386	0.364	0.966	0.862	1.142
6	20	143	0.363	0.386	0.364	0.966	0.862	1.142
6	23	186	0.363	0.386	0.382	1.017	0.838	1.132
6	26	210	0.363	0.386	0.382	1.078	0.804	1.124
6	29	266	0.363	0.386	0.379	1.119	0.795	1.122
6	32	261	0.363	0.386	0.368	1.119	0.795	1.122
6	35	310	0.363	0.375	0.376	1.192	0.781	1.066
6	38	294	0.363	0.386	0.376	1.176	0.781	1.098
6	41	448	0.363	0.386	0.377	1.263	0.769	1.030
6	44	329	0.363	0.386	0.374	1.219	0.781	1.066
6	47	324	0.363	0.386	0.374	1.219	0.781	1.066
6	50	361	0.363	0.386	0.375	1.233	0.771	1.066
6	53	371	0.363	0.386	0.378	1.243	0.771	1.057
6	56	408	0.363	0.386	0.379	1.265	0.769	1.041
6	59	520	0.363	0.386	0.373	1.307	0.749	1.026

Table 12—(Continued)

Ensemble	Target Percentile	Realization	k_{rw} end point region 1	k_{rw} end point region 2	k_{rw} end point region 3	Porosity Multiplier region 1	Porosity Multiplier region 2	Porosity Multiplier region 3
6	62	465	0.363	0.386	0.366	1.290	0.758	1.030
6	65	595	0.363	0.392	0.367	1.327	0.745	1.009
6	68	580	0.363	0.389	0.373	1.327	0.746	1.011
6	71	567	0.363	0.389	0.372	1.327	0.746	1.011
6	74	621	0.363	0.392	0.373	1.332	0.745	1.009
6	77	618	0.363	0.392	0.369	1.332	0.745	1.009
6	80	627	0.363	0.392	0.374	1.332	0.745	1.009
6	83	556	0.363	0.389	0.367	1.328	0.749	1.011
6	86	730	0.362	0.392	0.366	1.344	0.738	1.006
6	89	678	0.363	0.392	0.368	1.346	0.741	1.006
6	92	679	0.363	0.392	0.368	1.346	0.741	1.006
6	95	675	0.363	0.392	0.368	1.346	0.741	1.009
6	98	653	0.363	0.392	0.368	1.346	0.741	1.009
7	2	85	0.363	0.386	0.370	0.887	0.908	1.142
7	4	88	0.363	0.386	0.370	0.887	0.908	1.142
7	6	6	0.365	0.386	0.364	0.823	0.958	1.101
7	8	113	0.363	0.386	0.368	0.934	0.871	1.142
7	10	82	0.363	0.386	0.370	0.887	0.928	1.142
7	12	101	0.363	0.386	0.374	0.919	0.897	1.142
7	14	157	0.363	0.386	0.374	0.982	0.840	1.142

Table 12—(Continued)

Ensemble	Target Percentile	Realization	k_rw end point region 1	k_rw end point region 2	k_rw end point region 3	Porosity Multiplier region 1	Porosity Multiplier region 2	Porosity Multiplier region 3
7	16	161	0.363	0.386	0.374	0.982	0.839	1.142
7	18	191	0.363	0.386	0.382	1.032	0.838	1.124
7	20	98	0.363	0.386	0.374	0.919	0.908	1.142
7	22	187	0.363	0.386	0.382	1.017	0.838	1.132
7	24	190	0.363	0.386	0.382	1.032	0.838	1.124
7	26	211	0.363	0.386	0.382	1.078	0.804	1.125
7	28	263	0.363	0.386	0.368	1.119	0.795	1.122
7	30	222	0.363	0.386	0.382	1.100	0.804	1.125
7	32	264	0.363	0.386	0.379	1.119	0.795	1.122
7	34	255	0.363	0.386	0.368	1.119	0.795	1.125
7	36	241	0.363	0.386	0.379	1.110	0.811	1.125
7	38	444	0.363	0.386	0.377	1.263	0.769	1.031
7	40	441	0.363	0.386	0.377	1.263	0.769	1.031
7	42	294	0.363	0.386	0.376	1.176	0.781	1.098
7	44	340	0.363	0.386	0.375	1.219	0.781	1.066
7	46	318	0.363	0.386	0.374	1.219	0.781	1.066
7	48	362	0.363	0.386	0.375	1.233	0.771	1.064
7	50	370	0.363	0.386	0.379	1.243	0.771	1.057
7	52	416	0.363	0.386	0.378	1.265	0.769	1.032
7	54	516	0.363	0.386	0.373	1.307	0.749	1.026
7	56	354	0.363	0.386	0.375	1.233	0.777	1.066
7	58	467	0.363	0.386	0.366	1.290	0.758	1.030

Table 12—(Continued)

Ensemble	Target Percentile	Realization	k_{rw} end point region 1	k_{rw} end point region 2	k_{rw} end point region 3	Porosity Multiplier region 1	Porosity Multiplier region 2	Porosity Multiplier region 3
7	60	465	0.363	0.386	0.366	1.290	0.758	1.030
7	62	514	0.363	0.386	0.373	1.307	0.749	1.026
7	64	407	0.363	0.386	0.379	1.265	0.769	1.043
7	66	520	0.363	0.386	0.373	1.307	0.749	1.026
7	68	524	0.363	0.386	0.371	1.307	0.749	1.026
7	70	575	0.363	0.389	0.372	1.327	0.746	1.011
7	72	618	0.363	0.392	0.369	1.332	0.745	1.009
7	74	511	0.363	0.386	0.373	1.307	0.749	1.026
7	76	603	0.363	0.392	0.367	1.332	0.745	1.009
7	78	625	0.363	0.392	0.374	1.332	0.745	1.009
7	80	620	0.363	0.392	0.373	1.332	0.745	1.009
7	82	615	0.363	0.392	0.369	1.332	0.745	1.009
7	84	637	0.363	0.392	0.373	1.334	0.745	1.009
7	86	726	0.362	0.392	0.366	1.344	0.738	1.006
7	88	676	0.363	0.392	0.368	1.346	0.741	1.006
7	90	654	0.363	0.392	0.368	1.346	0.741	1.009
7	92	686	0.363	0.392	0.368	1.346	0.741	1.006
7	94	733	0.360	0.392	0.368	1.344	0.738	1.006
7	96	549	0.363	0.389	0.367	1.328	0.749	1.013
7	98	667	0.363	0.392	0.368	1.346	0.741	1.009

5. PRODUCTION OPTIMIZATION

5.1 Introduction

In this chapter, we investigate the production optimization strategies to be applied to the set of selected models in the previous chapter. Here, the optimized controls are yearly production and injection rate for each well during the next ten years. Controls for each ensemble are optimized to maximize expected net present value. Due to the significant number of control variables, direct perturbation is not practical because it requires a significant number of runs. Simultaneous perturbation stochastic approximation (SPSA) is used in this study because the number of runs per iteration in SPSA is not dependent on the number of control variables.

5.2 Assumptions and Constrains

The following assumptions and constraints are made in this study. The value of assumptions and constraint used in this study are commonly used in oil and gas industry:

- Oil price: 85$/STB

- Water processing cost: 15$/STB

- Water injection cost:15$/STB

- Discount rate: 10% yearly

- Total injection and production at every time step are equal

- Maximum water injection rate/well: 5000 BBL/D

- Maximum liquid production rate/well: 5000 BBL/D

5.3 Objective Function

The objective function for this optimization study is the expected NPV of the ensemble. The objective function used in this study is based on the objective function given by Brouwer and Jansen (2004), which can be written as:

$$J(q_{1:K}) = \sum_{k=1}^{K} J_K(q_{1:k})$$

$$= \sum_{k=1}^{K} \frac{\Delta t_k \left[r_o * q_{o,k}(q_{1:K}) - r_{wp} * q_{wp,k}(q_{1:k}) - r_{wi} * q_{wi,k} \right]}{(1 + b_\tau)^{\frac{t_k}{\tau}}} \tag{20}$$

where:

$J(q_{1:K})$ is net present value

q_k is vector of all phases flow rate

$q_{o,k}$ is vector of oil production rate

$q_{wp,k}$ is vector of water production rate

$q_{wi,k}$ is vector of water injection rate

r_o is oil price

r_{wp} is water processing cost

r_{wi} is water injection cost

b_τ is discount rate

Δt_k is time step size

K is total number of time step

t_k is cumulative time until k

The objective function in Eq. (20) accounts for only a single model. Here the interest is in optimizing for the entire ensemble. Thus, Eq. (20) can be rewritten to account for all models in the ensemble which is the robust objective function:

$$J_{rob}(q_{1:K}) = \sum_{i=1}^{N} W_i * J_i(q_{1:K}) \tag{21}$$

where:

J_{rob} is robust objective function

N is number of model in ensemble

W_i is weighting factor for each realization

Since each model in the ensemble is independent, the gradient of the objective function in Eq. (21) is simply the weighted summation of each model's gradient in the ensemble.

$$\frac{\partial J_{rob}(q_{1:K})}{\partial q_k} = \sum_{i=1}^{N} W_i * \frac{\partial J_i(q_{1:K})}{\partial q_k} \tag{22}$$

The gradient of each model is approximated using SPSA (see Section 1.2.4.1). Once optimization process is completed, the optimum control for each ensemble is obtained. Then, each control path is applied back to the whole set of possible realizations (773 models) to obtain the distribution of NPV based on each control strategy.

5.4 Optimization Result

5.4.1 NPV and Optimum Control Path

After implementing SPSA to each ensemble and the MAP model (see Section 3.3), the optimum control path for each ensemble can be obtained. The resulting NPV

distributions are shown in Fig. 36. The average and standard deviation from each case are shown in **Table 13**. The plot of average NPV vs. number of models is shown in Fig. 37. The production and injection rate for each well for every ensemble is shown in Fig. 38, Fig. 39 and Fig. 40. The description for each ensemble is as explained in Section 3.3. Each optimum control path is then input back into all accepted realizations from the original ensemble (773 models).

The result shows below illustrated the incremental NPV from using multiple models. The resulting NPV increased with number of models used during optimization process. 80% of the maximum incremental NPV can be achieved by using only 9 models. However, the incremental NPV diminish with increasing number of model in the ensemble. This is because as number of model in ensemble increase the remaining uncertainty space that has not been taking into account for optimization become less and less. Thus, the additional uncertainty space that each additional model is exploring becomes less significant. We do not observe any significant improvement in term of reducing risk of the net present value as there is no significant difference in standard deviation between each case.
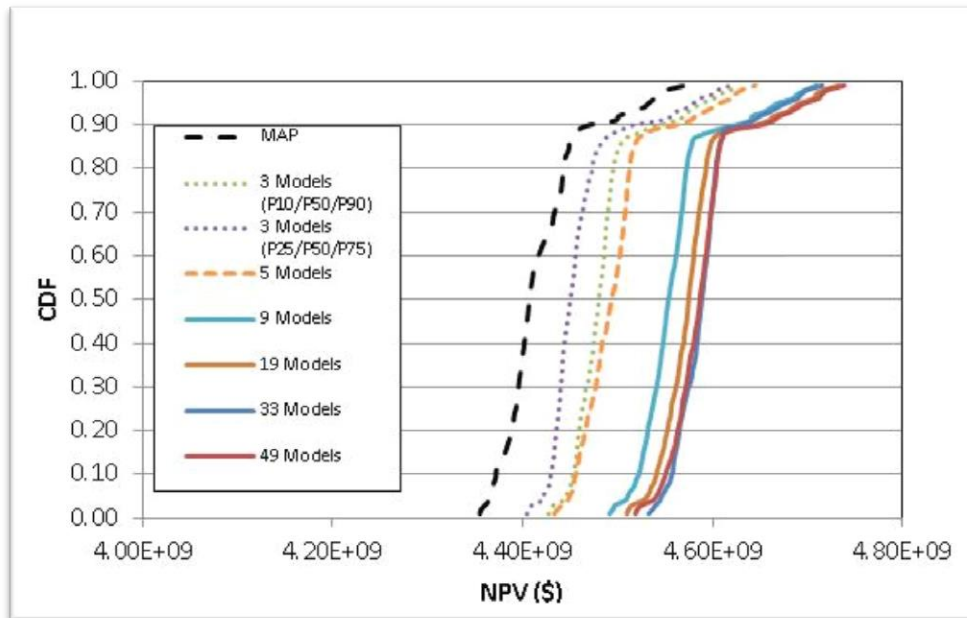
Fig. 36—Cumulative distribution functions of NPV of all realizations based on optimum control paths

Table 13—Summary of cumulative distributions of NPV based on control paths from different ensembles

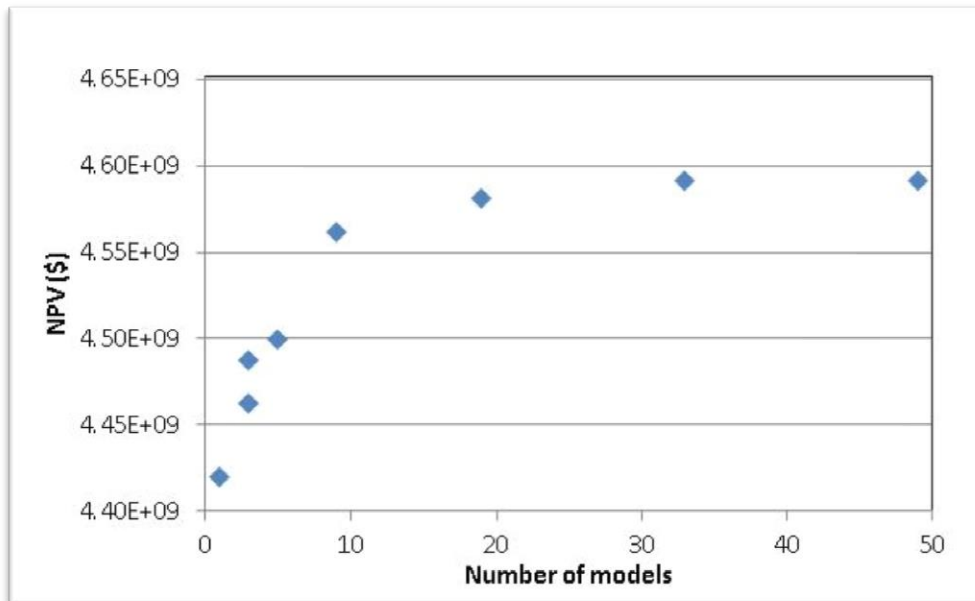| Scenario	Number of models	Average NPV ($)	NPV Standard deviation ($)
MAP	1	4.42E+09	4.59E+07
Ensemble 1	3	4.49E+09	4.08E+07
Ensemble 2	3	4.46E+09	4.46E+07
Ensemble 3	5	4.50E+09	4.39E+07
Ensemble 4	9	4.56E+09	4.39E+07
Ensemble 5	19	4.58E+09	4.41E+07
Ensemble 6	33	4.59E+09	3.53E+07
Ensemble 7	49	4.59E+09	4.28E+07

Fig. 37—Plot of average NPV using optimum control strategies from different ensemble sizes
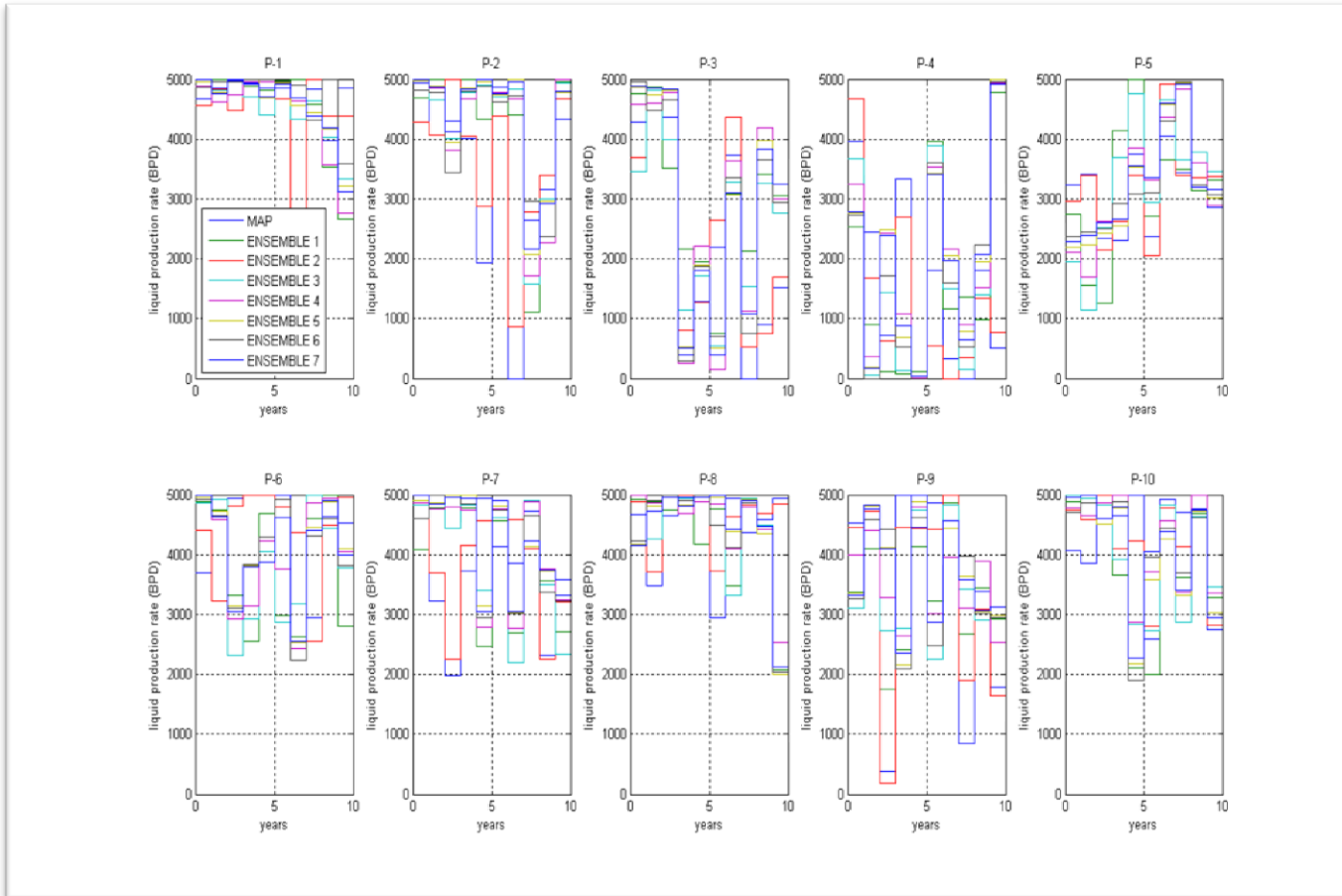
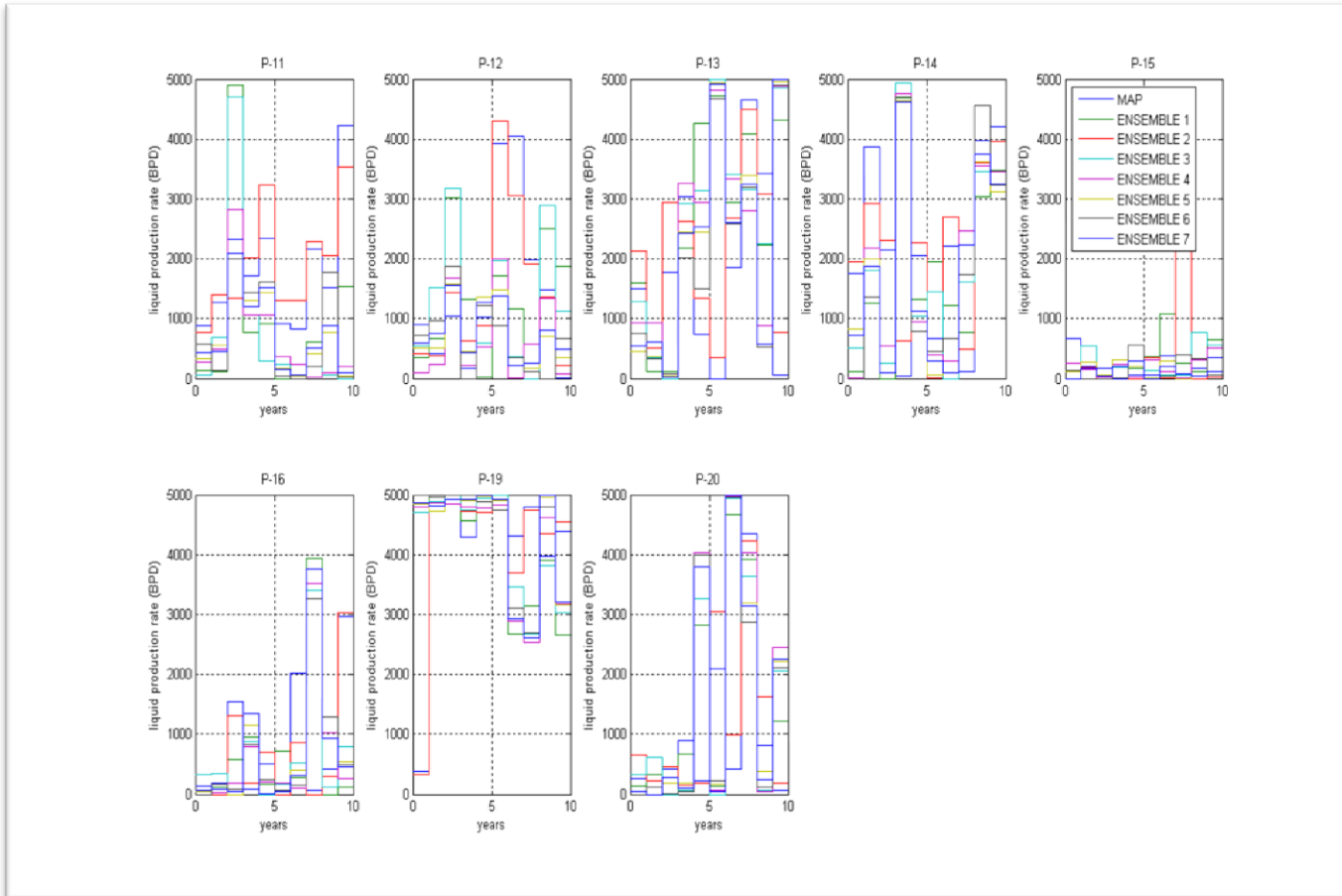**Fig. 38**—Optimal control strategies for wells P-1 through P-10

86

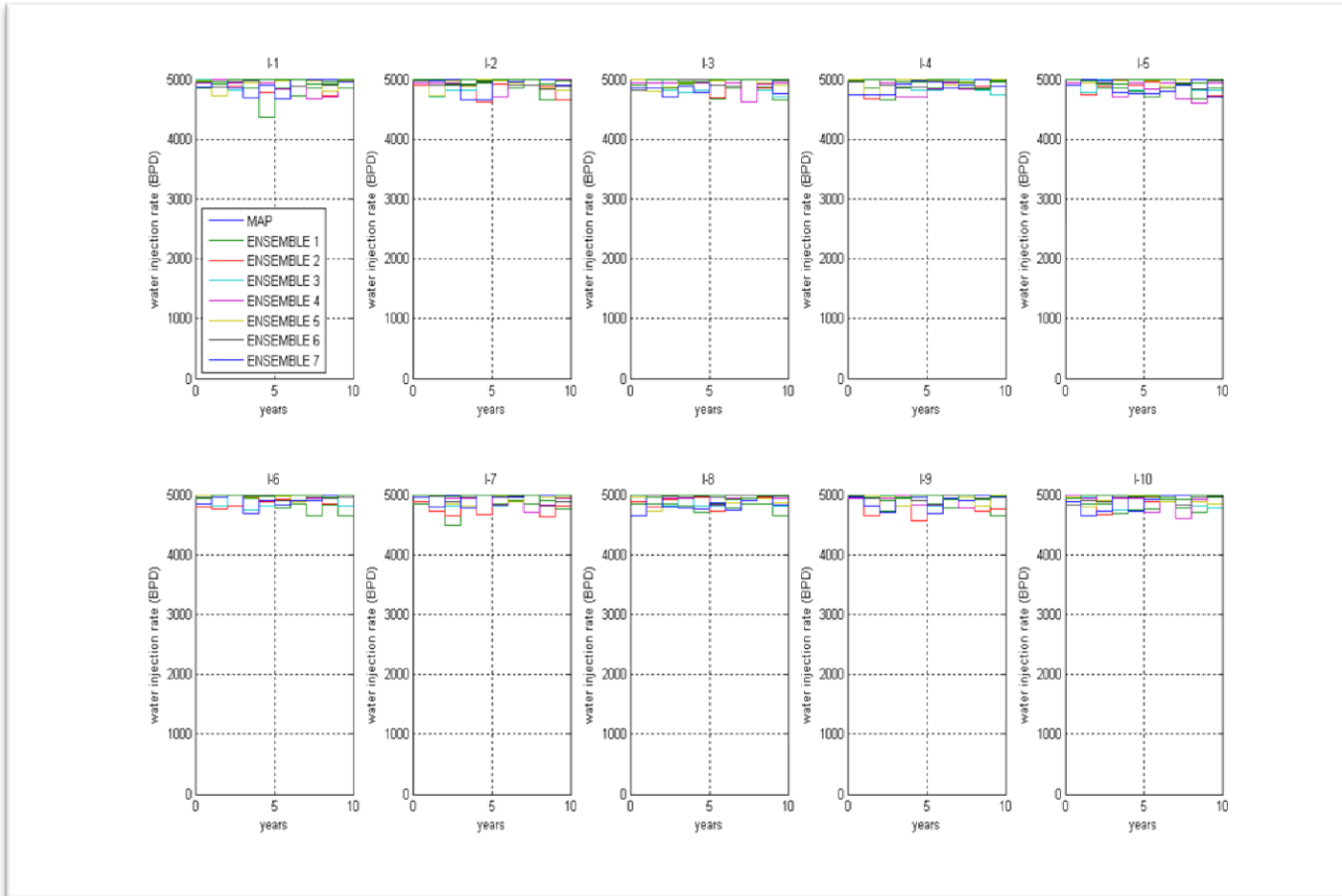**Fig. 39**—Optimal control strategies for wells P-11 through P-20

87

Fig. 40—Optimal control strategies for wells I-1 through I-10

5.4.2 Net Present Value of True Case

The optimum control strategies from each scenario are applied to true case (actual reservoir model) to determine the resulting NPV for each case (**Table 14**). We observed increasing trend of true case NPV vs. number of model used in optimization process. We believed this is because the higher the numbers of model used in optimization process, the more uncertainty is taken into account during optimization process.

Table 14—True case NPV based on control strategy

Control Scenario	NPV True case ($)
MAP	4.30E+09
Ensemble 1	4.34E+09
Ensemble 2	4.31E+09
Ensemble 3	4.34E+09
Ensemble 4	4.38E+09
Ensemble 5	4.40E+09
Ensemble 6	4.43E+09
Ensemble 7	4.41E+09

5.4.3 Production Forecasts from each Ensemble

Other benefit of using multiple models for optimization is improvement in the quality of the production forecast. With multiple realizations, distributions of production forecasts can be obtained. The water cut of the true case and the minimum, average and maximum water cut from each scenario are shown in Fig. 41 and Fig. 42. The minimum and maximum water cut are based on the overall water cut of the profile.

We observed that increasing number of realizations provide wider range of water cut forecast which can bracket the true case better.
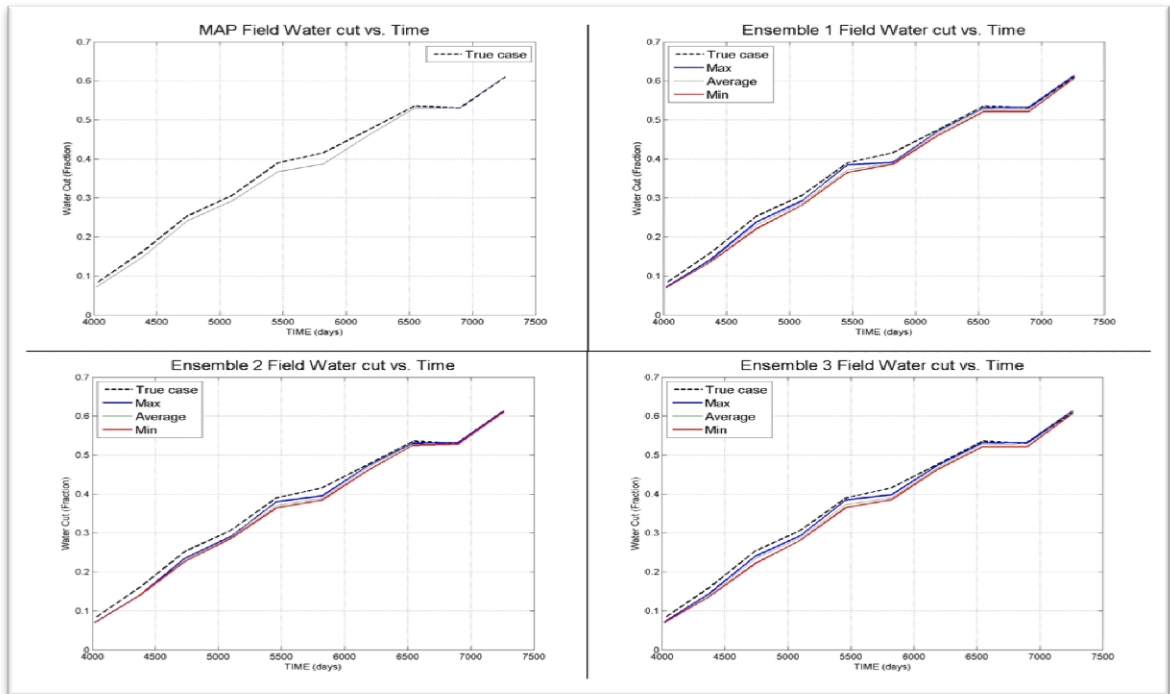

Fig. 41—Water cut forecast of MAP model and ensemble one through three compared to true case.
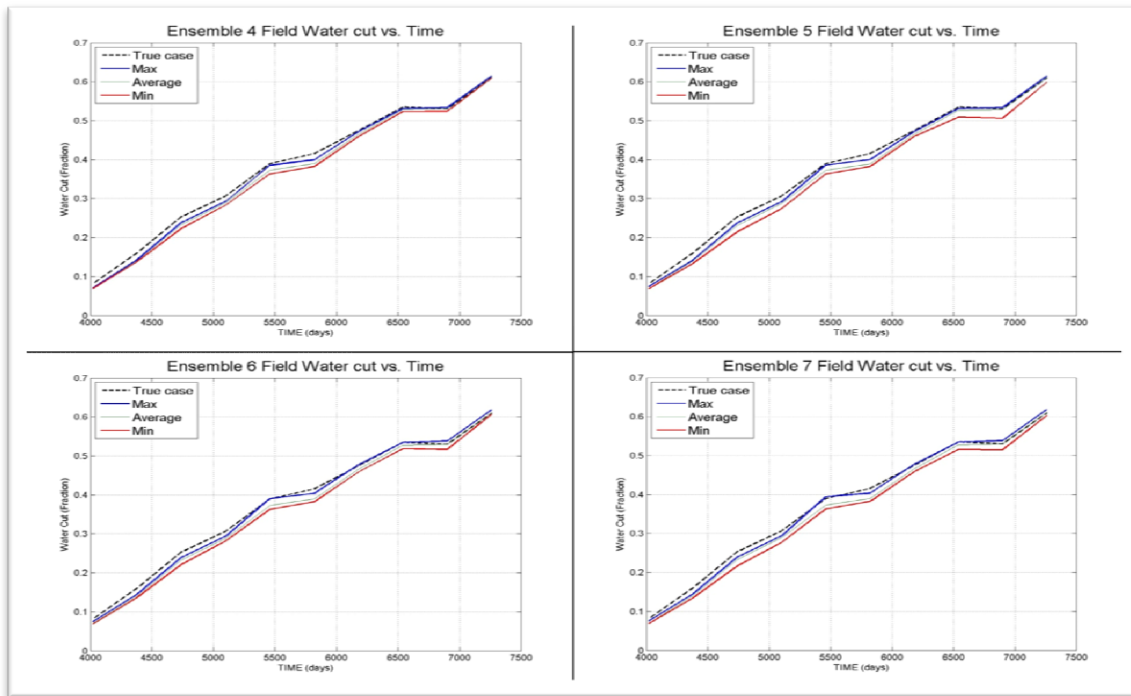
Fig. 42—Water cut forecast of MAP model and ensemble four through seven compared to true case.

5.5 Computational Cost

The results shown in Section 5.4.1 showed incremental NPV when additional realizations are used in the optimization process. However, including additional realizations into the optimization problem leads to significantly higher computational cost. The time required to run one flow simulation in this study is 45 seconds. Three runs per model are required for each iteration. Thus, the time required for one model in each of the iteration is approximately 2.5 minutes. Fig. 43 shows the computation time for different ensemble sizes for one thousand iterations. We observed that computation time increased linearly with numbers of models used in the optimization process as shown in Fig. 43. Fig. 37 indicated that there is no incremental benefit to increase number of

model more than 33 models as doing so will only increase computation cost without additional NPV.
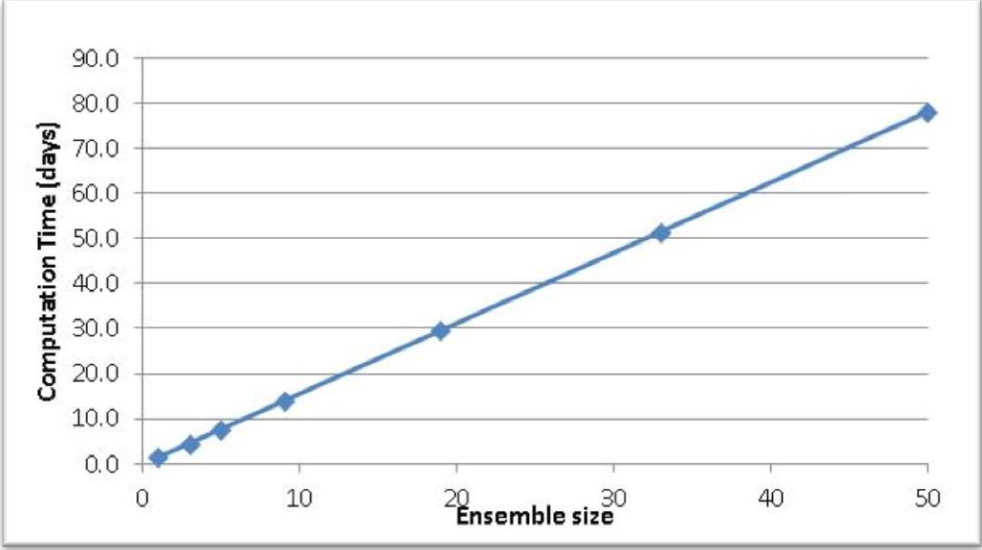

Fig. 43—Computation time for different ensemble sizes

6. CONCLUSIONS AND RECOMMENDATIONS

6.1 Conclusions

This study shows that there is benefit in using multiple models in the optimization process to account for uncertainty in reservoir parameters. The incremental NPV using multiple models observed in this study is up to 3.8% compared to using a single most-likely realization for optimization. However, the incremental NPV from additional models used in the optimization process diminishes with increasing numbers of model realizations. By using nine models, we can achieved 82% of the maximum benefit. I did not observe additional benefit to increasing the number of models to more than about 33 models. However, the computation cost increase linearly with increasing number of models use in optimization process. Thus, increasing number of model to be more than 33 models will only leads to additional computation cost without increasing NPV.

6.2 Assumptions, Limitations and Recommendations for Future Work

This study is based on the assumption that our history matching workflow is consistent to producing an ensemble of possible realizations that represent reservoir uncertainty reliably. If the ensemble of models does not represent well the level of uncertainty that exists in the reservoir, the optimum control path from any ensemble size may not be able to maximize NPV, because the basis for optimization is not correct.

There are several areas that require further study to improve the concept of using multiple realizations for optimization:

1. Incorporate the concept of hierarchical optimization by adding a second objective function to minimize the standard deviation within the ensemble. This should lead to optimum control paths that not only maximize the expected NPV but also minimize the risk.

2. Apply different rankings parameter during model selection, e.g., time of flight. In this study, field and region OOIP are used as ranking parameters. However, in some cases OOIP may not have direct correlation with NPV, which will lead to ensembles that cannot fully explore the uncertainty space.

3. Combine the concept of optimizing with multiple models with real-time reservoir simulation to investigate the improvement in production forecast accuracy in real time.

4. Implement different zonation strategies between true case and during history matching process to mimic case that we cannot correctly perform reservoir parameterization. Then, compare multiple-models optimization between cases with incorrect zonation of uncertain parameters (high spread in uncertain parameters) and the case with correct zonation.

NOMENCLATURE

p(m|d)	posterior probability
p(d|m)	likelihood probability
p(m)	prior probability
p(d)	probability of observed data
$\hat{g}_k(\hat{\theta}_k)$	approximated gradient
$\hat{\theta}_k$	vector of solution at current iteration
$\hat{\theta}_{k+1}$	vector of solution at next iteration
a_k	step size
GA	genetic algorithm
MAP	maximum a posteriori estimation
SPSA	simultaneous perturbation stochastic approximation
k	number of iteration
a	positive coefficient
A	positive coefficient
α	positive coefficient
c	positive coefficient
γ	positive coefficient
C_k	vector of perturbation size
Δ_k	user-specified random perturbation vector
$y(\hat{\theta}_k + C_k\Delta_k)$	measurement value at $\hat{\theta}_k + C_k\Delta_k$

$y(\hat{\theta}_k - C_k\Delta_k)$ measurement value at $\hat{\theta}_k - C_k\Delta_k$

k_v	vertical permeability
k_h	horizontal permeability
k_{rw}	relative permeability to water
m_i	uncertain parameter vector at current state
m_{i+1}	uncertain parameter vector at proposed state
w_j	random variable with distribution independent of the chain
OOIP	original oil in place
m_o	uncertain parameter vector as per prior knowledge
C_M	prior covariance matrix
c	normalizing constant
d_{obs}	observed data
g(m)	production profile from flow simulation
C_D	likelihood covariance matrix
α	acceptance probability
σ_{prior}	prior standard deviation
σ_{noise}	prior noise
ε	tolerance from target percentile.
M	number or output parameter.
P	number of representative model.
y	vector of output parameter
x	vector of input parameter

■	set of large but finite model
$\hat{\blacksquare}$	set of statistically representative model
q_k	vector of all phases flow rate
$q_{o,k}$	vector of oil production rate
$q_{wp,k}$	vector of water production rate
$q_{wi,k}$	vector of water injection rate
r_o	oil price
r_{wp}	water processing cost
r_{wi}	water injection cost
b_τ	discount rate
Δt_k	time step size
K	total number of time step
t_k	cumulative time until k
J_{rob}	robust objective function
N	number of model in ensemble
W_i	weighting factor for each realization
NPV	net present value

REFERENCES

Alhuthali, A. H. H., Datta-Gupta, A., Yuen, B. B. W. & Fontanilla, J. P. 2008. Optimal Rate Control Under Geologic Uncertainty.SPE-113628-MS. *SPE/DOE Symposium on Improved Oil Recovery.*Tulsa, Oklahoma, USA,20-23 April 2008

Alpak, F. O. & Kats, F. V. 2009. Stochastic History Matching of a Deepwater Turbidite Reservoir.SPE-119030-MS. *SPE Reservoir Simulation Symposium.* The Woodlands, Texas,2-4 February 2009

Bi, Z., Oliver, D.S., Reynolds, A.C. 2000. Conditioning 3D Stochastic Channels to Pressure Data. *SPE Journal*. 5(4), 474–484. DOI: 10.2118/67954-PA

Bieker, H. P., Slupphaug, O., & Johansen, T. A. 2007. Real-Time Production Optimization of Oil and Gas Production Systems: A Technology Survey. *Society of Petroleum Engineers*. 22(4), 382-391. DOI: 10.2118/99446-PA

Brashear, J. P., Becker, A. B. & Faulder, D. D. 2001. Where Have All the Profits Gone? *Journal of Petroleum Technology,* 53 (6)**,** 20-23, 70-73. DOI: 10.2118/73141-ms

Capen, E. C. 1976. The Difficulty of Assessing Uncertainty (includes associated papers 6422 and 6423 and 6424 and 6425). *Journal of Petroleum Technology,* 28**,** 843-850.

Chen, W., Xie, J. & Sarma, P. 2013. Selecting Representative Models From a Large Set of Models.SPE-163671-MS. *2013 SPE Reservoir Simulation Symposium.* The Woodlands, TX, USA, 18 – 20 Feb 2013

Deutsch, C. V., & Srinivasan, S. 1996. Improved Reservoir Management Through Ranking Stochastic Reservoir Models. *Society of Petroleum Engineers*. DOI: 10.2118/35411-MS

Dossary, M. & Mcvay, D. A. 2012. The Value of Assessing Uncertainty.SPE-160189-MS. *SPE Annual Technical Conference and Exhibition*. San Antonio, Texas, USA, 8-10 October 2012

Essen, G. V., Zandvliet, M., Hof, P. V. D., Bosgra, O. & Jansen, J.-D. 2009. Robust Waterflooding Optimization of Multiple Geological Scenarios. *SPE Journal,* 14**,** 202-210.

Gao, G., Li, G., & Reynolds, A. C. 2007. A Stochastic Optimization Algorithm for Automatic History Matching. *Society of Petroleum Engineers*. DOI: 10.2118/90065-PA

Gildin, E., & Lopez, T. 2011. Closed-Loop Reservoir Management: Do we Need Complex Models? *Society of Petroleum Engineers*. DOI: 10.2118/144336-MS

Geweke, J. 1992. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments. *Bayesian Statistics 4*, ed. Bernardo, J.M.Berger, J.O.Dawid, A.P.et al.: Oxford University Press, Oxford, United Kingdom.

Gonzalez, R., Mcvay, D. & Fondren, M. 2013. Applying Calibration to Improve Uncertainty Assessment.SPE-166422-MS. *SPE Annual Technical Conference and Exhibition*.New Orleans, Louisiana, USA, 30 September- 2 October 2013

Hastings, W.K. 1970. Monte Carlo Sampling Methods Using Markov Chains and Their Applications. *Biometrika*, 57, 97-109.

Holmes, J., McVay, D. A. & Senel, O. 2007. A System for Continuous Reservoir Simulation Model Updating and Forecasting.SPE-107566-MS. *Digital Energy Conference and Exhibition*.Houston, Texas, U.S.A.,11-12 April 2007

Hdadou, H., & McVay, D. A. 2014. The Value of Assessing Uncertainty in Oil and Gas Portfolio Optimization. *Society of Petroleum Engineers*. DOI: 10.2118/169836-MS

Jansen, J.-D., Brouwer, R. & Douma, S. G. 2009. Closed Loop Reservoir Management.SPE-119098-MS. *SPE Reservoir Simulation Symposium*.The Woodlands, Texas,2-4 February 2009

Liu, C. & McVay, D. A. 2010. Continuous Reservoir-Simulation-Model Updating and Forecasting Improves Uncertainty Quantification. *SPE Reservoir Evaluation & Engineering,* 13**,** 626-637.

Liu, N. and Oliver, D.S. 2003. Evaluation of Monte Carlo Methods for Assessing Uncertainty. *SPE Journal* 8 (2): 188-195. DOI: 10.2118/84936-pa

Metropolis, N.; Rosenbluth, A.W.; Rosenbluth, M.N.; Teller, A.H.; Teller, E. 1953. "Equations of State Calculations by Fast Computing Machines". *Journal of Chemical Physics* 21 (6): 1087–1092.

Mohamed, L., Christie, M. A. & Demyanov, V. 2010. Comparison of Stochastic Sampling Algorithms for Uncertainty Quantification. *SPE Journal,* 15**,** 31-38.

Odai, L., & Ogbe, D. O. 2011. An Approach for Ranking Realizations to Characterize Reservoirs for Fluid Flow Simulation. *Society of Petroleum Engineers*. DOI: 10.2118/150738-MS

Oliver, D.S., Chuna, L.B., and Reynolds, A.C. 1997. Markov Chain Monte Carlo Methods for Conditioning a Permeability Field to Pressure Data. *Mathematical Geology,* 29 (1): 61-91.

Oliver, D. & Chen, Y. 2011. Recent Progress on Reservoir History Matching: a Review. *Computational Geosciences,* 15**,** 185-221.

Osterloh, W. T. 2008. Use of Multiple-Response Optimization to Assist Reservoir Simulation Probabilistic Forecasting and History Matching.SPE-116196-MS. *SPE Annual Technical Conference and Exhibition.*Denver, Colorado, USA, 21-24 September 2008

Pajonk, O., Schulze-Riegert, R., Krosche, M., Hassan, M. & Nwakile, M. M. 2011. Ensemble-Based Water Flooding Optimization Applied to Mature Fields.SPE-142621-MS. *SPE Middle East Oil and Gas Show and Conference.* Manama, Bahrain, 25-28 September 2011

Peters, E., Chen, Y., Leeuwenburgh, O. & Oliver, D. S. 2013. Extended Brugge Benchmark Case for History Matching and Water Flooding Optimization. *Computers & Geosciences*, 50, 16-24.

Rotondi, M., Nicotra, G., Godi, A., Contento, F.M., Blunt, M., Christie, M.: Hydrocarbon Production Forecast and Uncertainty Quantification: a Field Application. SPE-102135. *SPE Annual Technical Conference and Exhibition*, San Antonio, Texas, USA, 24-27 September 2006

Schaaf, T., Coureaud, B. & Labat, N. 2008. Using Experimental Designs, Assisted History Matching Tools and Bayesian Framework to get Probabilistic Production Forecasts.SPE-113498-MS. *Europec/EAGE Conference and Exhibition.* Rome, Italy, 9-12 June 2008

Spall, J.C. (1998),Implementation of the Simultaneous Perturbation Algorithm for Stochastic Optimization, *IEEE Transactions on Aerospace and Electronic Systems*, 34, 817-823

Welsh, M. B., Bratvold, R. B. & Begg, S. H. 2005. Cognitive Biases in the Petroleum Industry: Impact and Remediation.SPE-96423-MS. *SPE Annual Technical Conference and Exhibition*. Dallas, Texas, 9-12 October 2005