THE UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

CLUSTERING-BASED ROBUST OPTIMIZATION OF SMART WELLS IN NATURALLY FRACTURED RESERVOIRS

by
Zhe Liu

A thesis submitted in partial fulfillment of

the requirements for the degree of Master of Science

in the Discipline of Petroleum Engineering

The Graduate School

The University of Tulsa

2016

T H E U N I V E R S I T Y O F T U L S A

THE GRADUATE SCHOOL

CLUSTERING-BASED ROBUST OPTIMIZATION OF SMART WELLS IN NATURALLY FRACTURED RESERVOIRS

by
Zhe Liu

A THESIS

APPROVED FOR THE DISCIPLINE OF

PETROLEUM ENGINEERING

By Thesis Committee

_____, Chair
Rami Younis

_____
Fahim Forouzanfar

_____
Ram Mohan

ABSTRACT

Zhe Liu (Master of Science in Petroleum Engineering)

Clustering-Based Robust Optimization of Smart Wells in Naturally Fractured Reservoirs

Directed by Rami Younis

57 pp., Chapter 5: Summary, Conclusions and Recommendations

(270 words)

In the development of naturally fractured reservoirs, the existence of natural fractures induces severe fingering and breakthrough. To manage the flood and to improve the ultimate oil recovery, we propose a numerical workflow to generate optimal production schedules for smart wells, in which well segments can be controlled individually.

In the first part of this work, the embedded discretized fracture model has been used to model the flow through the naturally fractured reservoir. In addition, stochastic gradient approximation methods have been used to compute gradients for production optimization using optimal control theory. Although EnOpt is a high performance algorithm for robust optimization, when a large ensemble is used, it requires a large amount of ensemble run for a gradient estimation which is computationally inefficient. To improve the efficiency of robust optimization, a clustering method is proposed to group similar fracture models based on their model parameters and response similarity. Both hierarchical clustering and the k-means method have been chosen as the clustering methods to group fracture models.

In the second part of this thesis, a data assimilation framework has been proposed to reduce the uncertainty introduced by natural fractures. Synthetic observed production data is generated by the EDFM fracture model. Then the Ensemble Smoother-with Multiple Data Assimilation algorithm is selected as the history matching method to update fracture parameters for a Dual Porosity Dual Permeability (DPDK) model. The DPDK models, which have been constrained by observation data, are treated as input ensemble models for robust optimization. The optimal well controls obtained from robust optimization would be approximate optimal well controls for the real scenario.

ACKNOWLEDGEMENTS

I would like to take this opportunity to express gratitude to the FuRSST (Future Reservoir Simulation Systems Technology) for funding this project and my study. I would like to express my sincere gratitude to my academic advisor, Dr. Rami Younis, for giving me continuous support, guidance, and encouragement throughout my graduate study. His personal and professional attitude always encourage me when I feel confused in life. I feel very proud to be his student.

I would like to extend sincere thanks to Dr. Fahim Forouzanfar and Dr. Ram Mohan for being on my thesis committee. I also would like to offer my sincere thanks to Dr. Fahim Forouzanfar for his academic and personal guidance.

I appreciate the friendship of FuRsst friends, Yuchen Zhang, Guotong Ren, Jiamin Jiang, Yuanshan Zhang, Chen Fang, Yuxuan Jin, Soham Sheth and Giorgiy Lutidze, who offered encouragement, knowledge and experience. I would like to give very special thanks to my best Chinese friends in Tulsa for their continuous support and encouragement. With them, I was able to overcome all obstacles on the road to completing this work.

Finally, I would like to dedicate this thesis to my beloved family; and to my parents, Zhixiang Liu and Chunhong Yang, for their endless love, support, and always being my anchor whenever I feel sad.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Water flooding processes are commonly applied in the oil industry to increase oil production and enhance oil recovery. In a naturally fractured reservoir, the natural fractures serve as highly conductive flow paths for the reservoir fluid, which will result in the breakthrough of the waterfront. The breakthrough of water would not only cause the inefficient water flood (low sweep effect) but would also cause the increase of water disposal fees. Maintaining a stable waterfront should be an important consideration to increase oil production and decrease disposal water production in the development of naturally fractured reservoirs. So determining the optimal BHP for production wells as well as injecting rates for injection wells are critical steps in preparing the development plan for a naturally-fractured reservoir.

Production optimization is a challenging due to the high dimension and complexity of the search space and nonlinearity introduced by natural fractures. Also, the optimization is computationally very demanding as every objective function evaluation requires a reservoir simulation run. Moreover, it is challenging to determine the number, distribution and properties of natural fractures, they increase nonlinearity and uncertainty of the optimization problem. Therefore, an efficient, robust production optimization framework for naturally fractured reservoirs is desirable.

Brouwer and Jansen (2004) developed a systematic dynamic optimization approach based on optimal control theory and investigated the scope for optimization under purely

pressure-constrained and purely rate-constrained operating conditions. They used ICV settings as the control variables in the pressure-constrained scenario where the wellbore pressures were preselected, i.e., the wellbore pressures were fixed throughout the optimization process. Naus et al. (2006) proposed an operational strategy for commingled production with ICVs using Sequential Linear Programming (SLP) which is based on the simplex algorithm, where ICVs were modeled as a multiplication factor of PI's with values between 0 and 1.0. Van Essen et al. (2010) implemented a gradient-based optimization technique to optimize the ICVs of a field-scale reservoir. They proposed a dynamic grouping approach based on a visualization of the optimal ICVs which were obtained at the first stage of their workflow. However, as observed by Fonseca et al. (2015b), having fewer ICVs results in a loss of controllability and thus an optimized strategy with a lower NPV. Alhuthal et al. (2010) proposed an approach that relies on finite-difference or streamline-based models to optimize the production/injection rates of designated ICVs which maximizes the water flooding sweep efficiency, where the optimization was performed under operational and facility constraints using a sequential quadratic programming approach. Li and Zhu (2011) proposed a procedure of applying temperature distribution data as a feedback to operate ICVs to achieve an approximately uniform flow distribution and to increase oil flow rate and delay early water breakthrough.

EnOpt has become a popular method for the estimation of optimal well controls (e.g., water rates and bottom hole pressures) in the oil and gas reservoir engineering optimization community. Inspired by the Ensemble Kalman Filter (EnKF) method, it was first introduced by Lorentzen et al. (2006) and Nwaozo (2006). The standard formulation of the EnOpt algorithm which uses an ensemble of randomly perturbed control vectors to

approximate a gradient of the objective function was proposed by Chen et al. (2009).Thereafter, several modifications of the Chen et al. formula have been made for production optimization (Do and Reynolds, 2013; Fonseca et al., 2015b).

As we mentioned in the preview section, randomly distributed natural fractures introduce a great amount of uncertainty into our production optimization system. In order to reduce uncertainty and generate a reliable optimization scheme, the production history can be used to better constrain the reservoir model. This data assimilation process is widely known in the oil industry as history matching. Two classes of approaches are commonly used to match production history in earth science: 1) gradient-based history matching methods 2) gradient-free history matching methods.

A gradient-based history matching problem can be treated as a high dimensional optimization problem that treats the posterior probability function as the objective function. Although the adjoint method can improve the efficiency of the backward run compared with finite difference and the directional methods, it is difficult to derive the formulation of the adjoint method for different parameters. In the last decade, ensemble-based methods, which are gradient-free methods have been widely investigated and applied for data assimilation of flow problems associated with petroleum reservoir history matching. The ensemble Kalman filter, which was summarized by Oliver and Chen (2010) is the most popular for history matching applications due to its efficiency and ease of implementation. However, with EnKF, there exist potential inconsistencies between the updated model parameters and states, and the method often yields a relatively poor data match. The Ensemble Smoother with Multiple Data Assimilation (ES-MDA) (Emerick and Reynolds, 2012) attempts to correct this defect of EnKF. Furthermore, the ES-MDA makes the

matching of more flexible parameters possible. In this work, ES-MDA has been used as a history matching method to generate proper ensembles for the robust optimization algorithm.

1.1 Research Contribution

The generation of an optimal development plan for NFRs is a practical problem of timely interest. The first contribution of this work is to apply production optimization algorithms to NFRs. In this work, the EnOpt algorithm has been used to compute stochastic gradients for production optimization. The geological uncertainty introduced by natural fractures also has been considered in robust optimization. However, there are substantial uncertainty in fracture number, position, and properties (permeability, aperture, porosity). So the ensemble size should be very large to represent all possible underground cases. Robust optimization with large ensemble size is extraordinarily inefficient and computationally expensive. To improve the efficiency of the robust optimization algorithm, three clustering methods are proposed to reduce redundant ensemble runs. By using clustering algorithms to group different fields, we could significantly improve the efficiency of robust optimization without scarifying expected NPV for the optimal case.

The second contribution of this work is to propose a closed-loop reservoir management framework to develop NFRs. Closed-loop reservoir management has two major steps: production optimization and data assimilation. In practice it is challenging to select parameters for DFMs. So the dual porosity DP-DK model has been applied as a forward model to match the synthetic observation data, which was generated by Embedded Discrete Fracture Model (EDFM). The fracture permeability and pore volume field will

indicate the probability distribution of fracture field. Then an ensemble of these fields will be used as input models for robust optimization algorithm to generate optimal production rate.

1.2 Thesis Outline

There are 5 chapters and one appendix in this thesis. At the beginning of Chapter 2, we briefly introduce the EDFM, which has been used to generate NFR models. The steepest ascent method, EnOpt method and formulation of objective function are also illustrated, in this chapter. Chapter 3 presents details of robust optimization for NFRs; the k-means and hierarchical clustering methods are also introduced to reduce redundant ensembles in the robust optimization. Both model parameters and responses for the different models have been chosen as different features to group the fracture model. The computational results for regular robust optimization and clustered optimization are compared and analyzed in this chapter. Chapter 4 covers the framework of history matching and optimization for NFRs. After assimilating the production data, these fields have been input as ensembles for robust optimization. Then the robust optimal well controls have been compared with real optimal well controls are generated using the true field. In the computational results section, we compare the DP-DK fields with the true DFM field, which are used to generate synthetic observation data after assimilation.

CHAPTER 2

PRODUCTION OPTIMIZATION IN NFRS

Natural fractures serve as highly conductive flow paths for reservoir fluid, increasing the reservoir's heterogeneity and effective permeability significantly over that attributable to the rock matrix. To develop a naturally fractured reservoir without a water drive, the only mechanism to produce oil would be the expansion of reservoir fluids, which would cause small estimated ultimate oil recovery. On the other hand, water injection would result in easy breakthrough of water through the fracture. This lead to an inefficient sweep as well as to significant water disposal costs. Due to the geological complexity and heterogeneity introduced by natural fracture, it is unrealistic to design a generic development strategy fit for any naturally fractured reservoirs.

In this chapter, a numerical framework has been proposed to simulate fluid flow and optimize well controls in naturally fractured reservoirs. To simulate reservoir fluid flow in the natural fractures, the embedded discretize fracture model (EDFM) is applied. To maximize cumulative oil production rate, while minimizing water production and injection rate, we define the NPV of production from a reservoir as the objective function to be maximized in the production optimization problem. We briefly explain the steepest ascent algorithm and the scaling scheme that have been used as a basic line search algorithm to find optimal BHPs for producers and injection rates for injectors. EnOpt is also introduced as a method to generate the stochastic gradient for the steepest ascent algorithm. In the computational results section, two examples are presented. In the first

example, the workflow to set up nominal optimization in NFRs was illustrated. In the second example, several optimization schemes with stochastic gradients are compared with each other. The detailed form of the specific production optimization problems defined in naturally fractured reservoirs is given in the following chapter.

2.1 Models and Methods

2.1.1 Embedded Discrete Fracture Method

In naturally fractured reservoirs, the randomly distributed fracture position, orientation and properties introduce a large amount of uncertainty into optimization system. J Jiang et al (2015) compared the performance of EDFM model with Unstructured Discrete Fracture Method (USDFM) and conclude that EDFM can provide better computational efficiency relative to USDFM while adequately capturing sharp transient in naturally fracture.

In EDFM, there are three kinds of noneighboring connections (NNCs), which are fracture-matrix connections, fracture-fracture intersections and fracture-well connections respectively. The formulation of these three NNCs is given as follows:

For a NNC between matrix and fracture cell, the transmissibility is

$$T_{mf} = \frac{k_{mf}A}{d_{mf}}, \tag{1}$$

where,

$$d_{mf} = \frac{\int x_n dv}{V}, \tag{2}$$

A is the fracture surface area in the grid block, and k_{mf} is defined as a harmonic average of the matrix and fracture permeability. Li and Lee (2008) assumed that pressure varies

linearly in the direction normal to each fracture within each grid block and using Eq.(2), the average normal distance ,d_{mf}, is computed for matrix fracture connections. The enumerated method and numerical integration have been used in our preprocessing code to calculate these parameters, where dv is matrix volume element, x_n is normal distance of the element from the fracture plane and V is volume of a grid block.

For the NNC between fracture intersections, we calculate the harmonic average of transmissibility,

$$T_{ff} = \frac{k_{ff}A}{d_{ff}} = \frac{2\,T_1T_2}{T_1+T_2}, \tag{3}$$

where,

$$T_i = \frac{k_{fi}w_{fi}L_{int}}{d_{fi}}\ (i = 1,2), \tag{4}$$

w_{fi} is aperture of the i^{th} fracture segment, k_{fi} is fracture permeability of the i^{th}fracture segment; $L_{int}$is the length of the intersection, and d_{fi} represents weighted distance from the i^{th} intersection to each fracture segment.

The well index for well bores intersected by natural fractures is

$$WI = \frac{2\pi wk_f}{\log(r_o/r_w)}, \tag{5}$$

where,

$$r_o = 0.14\sqrt{l^2 + h_f^2}, \tag{6}$$

l is the length of fracture segment and h_f is the fracture height.

2.1.2 Definition of NPV and Optimization Problem

The objective function in this work is the NPV of a time series of cash flow. For the two-phase flow of oil water, NPV is defined by:

$$NPV = \sum_{n=1}^{N}\{[\sum_{j=1}^{p}[q_{oj}n * r_0 - c_w * q_{wj}n] - \sum_{j=1}^{J}[c_{wi} * q_{iwj}n]] * \Delta t / 1 + b^{t/365}\}, \quad (7)$$

So the optimization problem is

$$\text{MAX } \{\text{NPV}\} \qquad \text{s.t. } u_{low} < u < u_{up}\,, \quad (8)$$

where r_0 is oil rate, c_w is the disposal cost of produced water, c_{wi} is the water injection cost, b is annual inflation rate. t_n is time at the end of the n^{th} period, Δt is n^{th} time step size, and N is the total number of time steps, P is the number of producer, I is the number of injector, $q_{oj}n$ denote the average oil rate and rate at the j^{th} producer well for the n^{th} time step, $q_{wj}n$ is the average water rate and rate at the j^{th} producer well for the n^{th} time step, whereas $q_{iwj}n$ is the average injecting rate at j^{th} injector well for the n^{th} time step. Note that in Eq.7 we have assumed constant economical parameters for whole production life of a reservoir.

2.1.3 Gradient Calculation

EnOpt is a stochastic gradient-based optimization method that uses an ensemble of control vectors to estimate an approximate gradient of the objective function with respect to the well control vector. A single control vector is defined as

$$u = [u_1, u_2, u_3, \dots, u_N], \quad (8)$$

Where *N* is a total number of control variables (e.g. BHP for producers, injecting rate for injectors over each control time). In the original EnOpt, a multivariate, Gaussian-distributed ensemble is generated with a distribution mean $\bar{u}$, ensemble size M, and a

predefined distribution covariance matrix. Several modifications of EnOpt formula have been made for production optimization (Do and Reynolds, 2013; Fonseca et al., 2015b). Instead of using $\bar{u}$, the control variable from last iteration step u_{j-1} is used. To estimate the gradient, a mean-shifted ensemble matrix is defined as:

$$\bar{U} = [u_1 - \bar{u}, u_2 - \bar{u}, \dots u_M - \bar{u}], \tag{9}$$

Similarly, a J -shifted objective function vector is defined as (where J is the value of objective function last timestep):

$$J = [J_1 - \bar{J}, J_2 - \bar{J} \dots J_M - \bar{J}], \tag{10}$$

The approximate gradient as proposed by the mean-shifted and J-shift is given by:

$$d_k = \frac{1}{N_e - 1} C_U \sum_{j=1}^{N_e} \left(\hat{u}^j - \overline{\hat{u}}\right)\left(J\left(\hat{u}^j\right) - \bar{J}\right) = \frac{1}{N_e - 1} C_U C_{Uf}, \tag{11}$$

where C_U is the covariance matrix used to smooth the production rate and can be obtained by using the covariance function. In the real field, controls of the certain well are correlated in time. The formulation of the temporal covariance matrix and covariance function are illustrated in Appendix A.

2.1.4 Steepest Ascent and Scaling of Variables

Steepest ascent is a first-order optimization algorithm. To find a local maximum of a function using gradient ascent, a step proportional to the gradient direction (or of the approximate gradient) of the function at the current point is taken. The updates for gradient ascent are

$$u_k = u_{k-1} + a_k d_k, \tag{11}$$

where u_k denotes the well control variable for k^{th} iteration. a_k is the step size at k^{th} iteration, d_k is the search direction at k^{th} iteration. The step size is modified by a simple

backtracking algorithm and the search direction is a stochastic gradient, which is generated by the EnOpt method.

The steepest ascent method is sensitive to the scaling of the problem. In production optimization problems, BHPs for producers and injection rates for injectors are not similar in magnitude. Hence this is ill-scaled problem, and the steepest ascent would have poor convergence. In order to improve the convergence rate, we rescale the problem:

$$0 \leq \tilde{u}_i = \frac{u_i - u_i^{\text{low}}}{(u_i^{\text{up}} - u_i^{\text{low}})} \leq 1, \tag{12}$$

where the index i indicate i^{th} well and u_i represent control vector for ith well. If the i^{th} well is a producer, u_i^{low} and u_i^{up} represent maximum and minimum BHP, respectively, for that well. And if the i^{th} well is an injector, u_i^{low} and u_i^{up} represent maximum and minimum injecting rates of the well, respectively.

The convergence criteria for the steepest ascent problem are set as if both objective function and control variables do not have significant changes. The formulation of the convergent criteria are:

$$\frac{|J(u^{k+1}) - J(u^k)|}{|J(u^k)|} \leq 10^{-4}\,, \tag{13}$$

and

$$\frac{\|u^{k+1} - u^k\|_2}{\max\left(\|u^k\|_2, 1.0\right)} \leq 10^{-3}, \tag{14}$$

2.2 Computational Results

2.1.1 Example 1: Nominal Optimization for NFRs

The first example is a two-dimension, two-phase (oil-water) homogenous NFR. The reservoir size is 2500ft and a grid of dimension $50 \times 50 \times 1$ is used. The reservoir permeability is 10 md and porosity is 0.2. There are two smart horizontal wells drilled parallel to each other. One is a producer and the other is an injector. The heel to toe coordinates of the production well are (125, 250) and (125, 2250) respectively. And the start and end points of the injection well are (2325, 250) and (2325, 2250). Each well is separated into 8 equal segments and each segment independently controlled by an ICV independently. The natural fractures are randomly distributed within the reservoirs as shown in Fig.2.2.1. The parameters used to generate the true fracture field are listed in Table 2.1. The oil price used is \$80/STB, disposal water costs is \$2/STB and injection water price is \$2/STB.

The steepest ascent algorithm using a stochastic gradient obtained from the EnOpt algorithm is chosen as the optimization scheme. The maximum bottom hole pressure (BHP) is set as the initial reservoir pressure (4000psi). Moreover, the minimum BHP is set at 1500 psi. As for the injection well, the maximum and minimum injection rates are set at 500 STB/d and 0 STB/d respectively.

	Mean Value	Standard Deviation	Units
Number of Fracture	15	3	
Fracture Length	1200	50	ft
Fracture Orientation	0	0.392	rad
Fracture Permeability	5000	200	md
Fracture Aperture	0.1	0.01	ft
Fracture Porosity	0.95	0.1	

Tab 2.1 The Basic Parameters to Generate Fracture Field

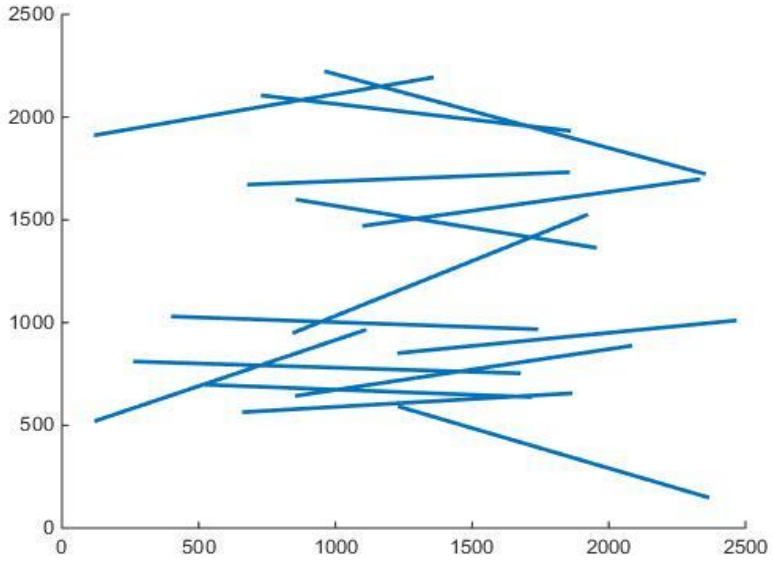

Fig 2.2.1 Model Sketch and Fractures Distribution

Number			

of Ensemble	10	**Maximum step**	0.1
Number of Well	16	**Pi**	3000
Number of Producer	8	**Swi**	0.15
Number Control Steps	30	**k**	10
Minimum Inj Rate	0	**phi**	0.2
Maximum Inj Rate	500	**Oil Price**	100
Minimum Pro Pressure	1500	**Disposals Water Rate**	5
Maximum Pro Pressure	3000	**Injecting Water Rate**	1
Maximum Iteration Steps	100	**Inflating Rate**	0.05

Tab 2.2 The Basic Parameter to Set up Optimization Algorithm

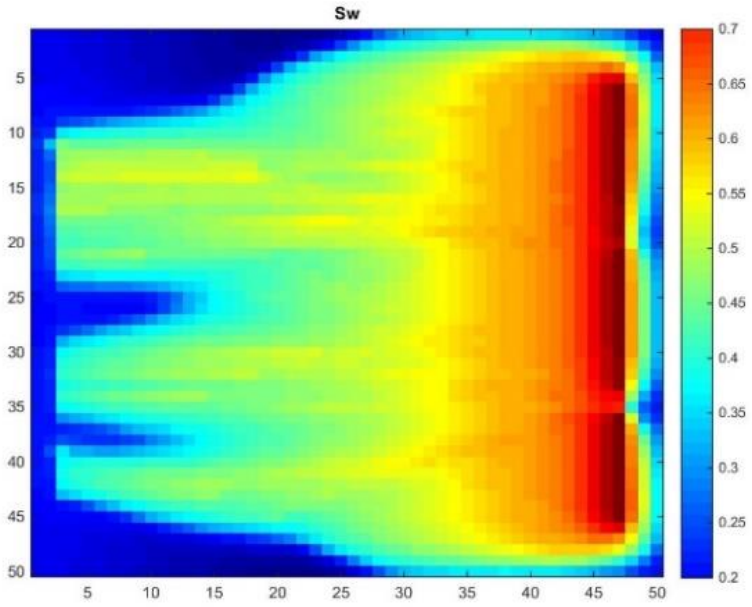

Fig 2.2.2 The Final Water Saturation Profile for Reference Case

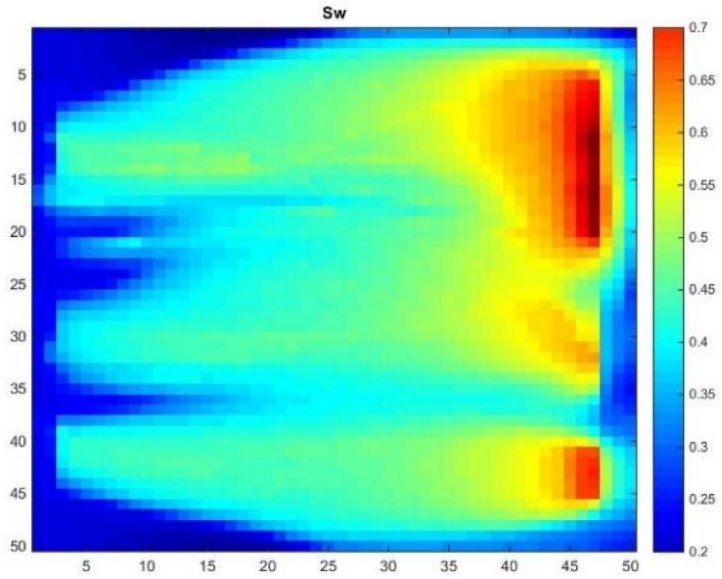

Fig 2.2.3 The Final Water Saturation Profile for Optimal Case

By comparing the final water saturation profile of the reference and optimal cases, it is observed that, in the optimal case, the water is more uniformly distributed in the optimal case than in the original case. This indicates that production using the optimal well controls could help to enhance sweep efficiency. Figs 2.2.4 and 2.2.5 demonstrate the optimal well controls of each well segment in each time step. At the end of production, the ICV for well segments 3 and 6 tend to close. By combining Figs 2.2.4 and 2.25 with 2.2.1, we could observe that the well segments tend to close where natural fractures are well developed. Fig2.2.4 indicates that injection well tends to inject water at the beginning with a large amount of water injected in well segments 1, 5 and 8. Figs 2.2.1 and 2.2.5 show that injecting water using an optimal well control scheme will avoid natural fractures.

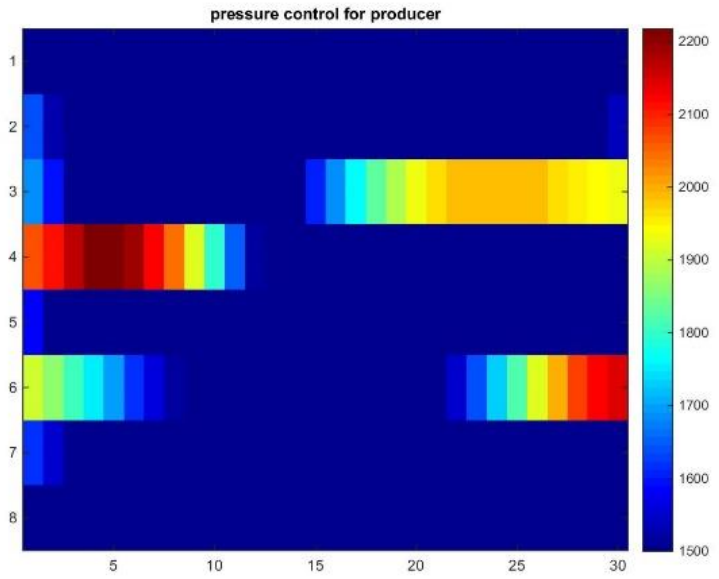

Fig 2.2.4 The Optimal BHP Controls for Production Well Segment (EnOpt Gradient)

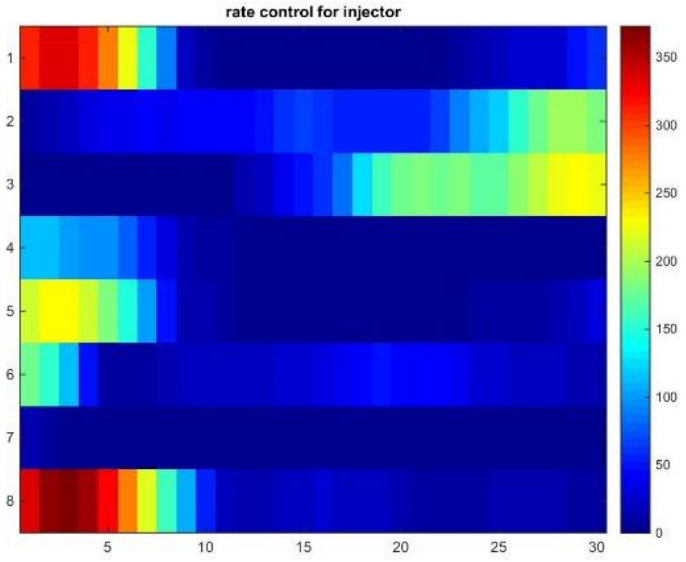

Fig 2.2.5 The Optimal Injecting Control for Injection Well Segment (EnOpt Gradient)

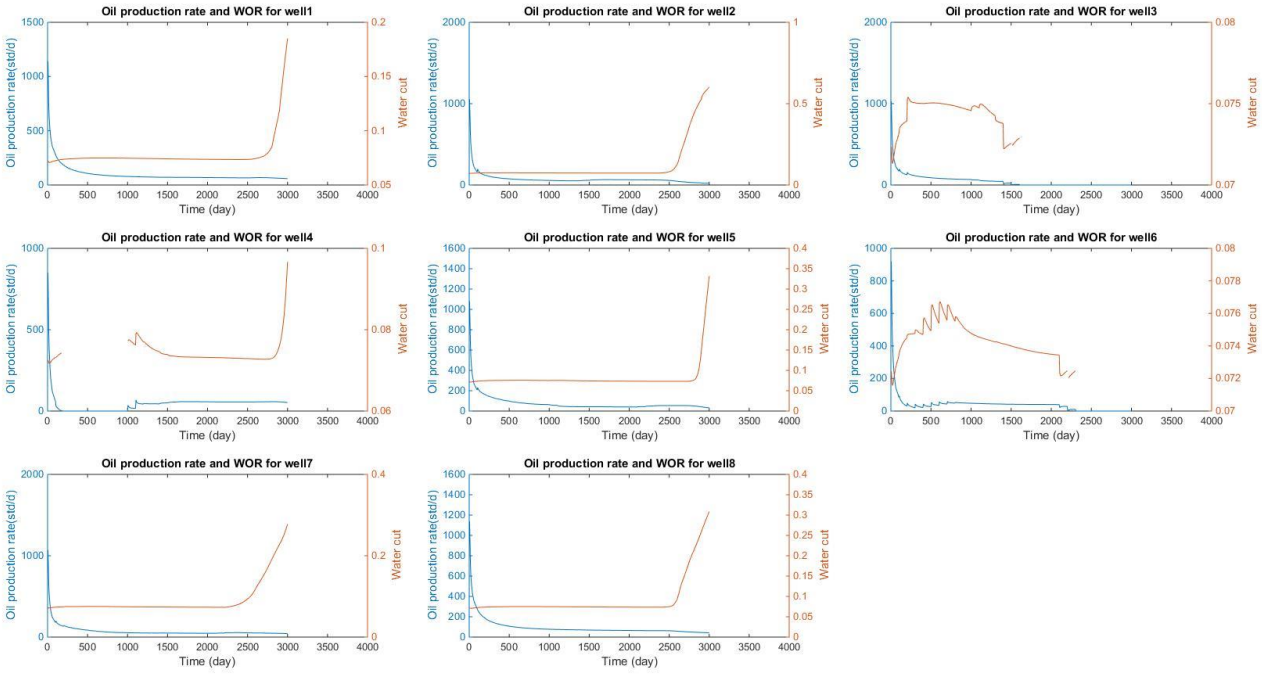

Fig 2.2.6.Oil Production Rate and Water Cut for Each Production Well Segments

2.1.2 Example 2: Comparison of Optimization Schemes

Figs 2.2.7 and 2.2.8 show optimal well controls for each well segment generated by the Simplex gradient. By comparing Fig-5 and Fig-6, we can see that both the EnOpt and the Simplex method lead to similar results. As to the producer, optimal control suggests that all segments start fully open and then shut in segment 3 at the end of production. The Simplex method has better performance than the EnOpt method. When it comes to the change in NPV, compared with the EnOpt gradient, the Simplex gradient has a better convergence rate and higher optimal NPV.

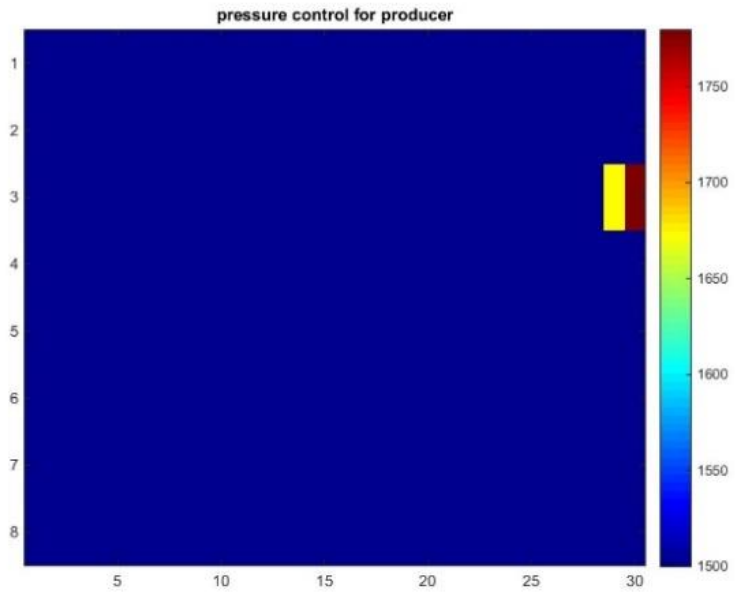

Fig2.2.7 The Optimal BHP Controls for Production Well Segment (Simplex Gradient)

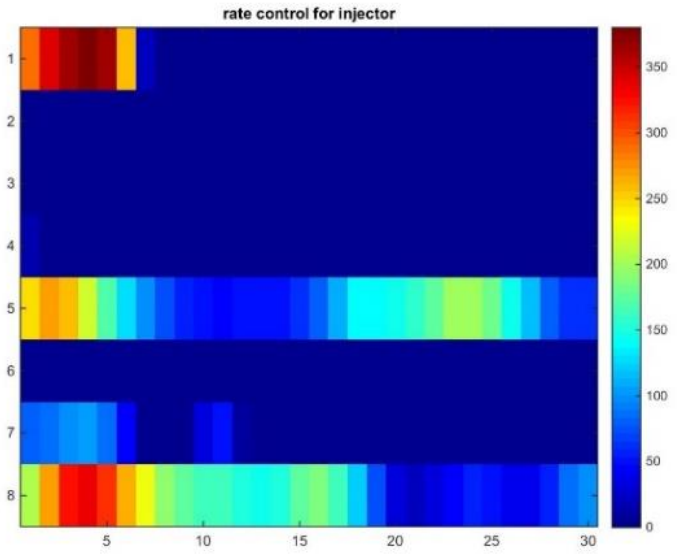

Fig 2.2.8 The Optimal Injecting Rate Control for Injection Well Segment (EnOpt Gradient)

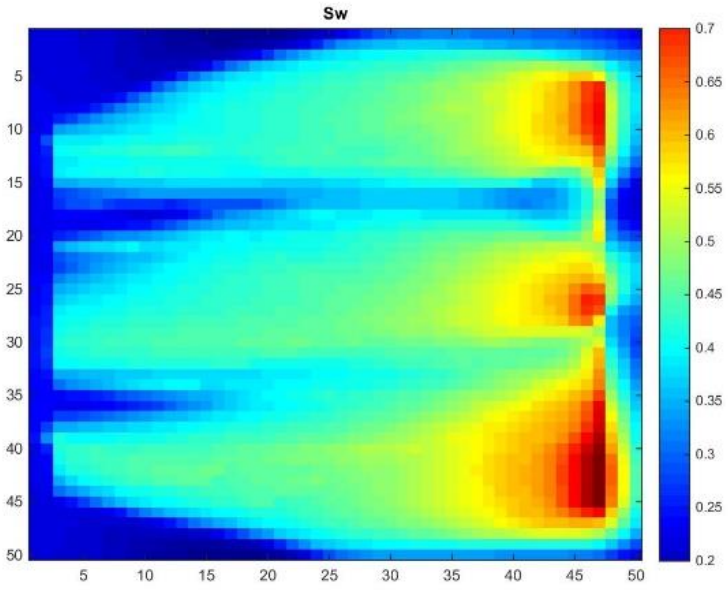

Fig.2.2.9. The Final Water Saturation Profile (Simplex Gradient)

Both the EnOpt and the Simplex method can enhance the cumulative oil production rate by 26% compared with the reference case. The optimal scheme generated by the EnOpt algorithm could produce more oil than other methods generated by Simplex gradient. However, when it comes to waterfront management, the Simplex method has better performance than EnOpt (as seen in Fig-8). In the optimal case generated by the EnOpt method, the waterfront reaches the production well at around 2600 days after the production. Using the Simplex method, the water does not break through to the producer within the reservoir lifetime (3000 days). Compared with the reference case, the steepest ascent method with Simplex gradient decreases water production of the production well by 32%.

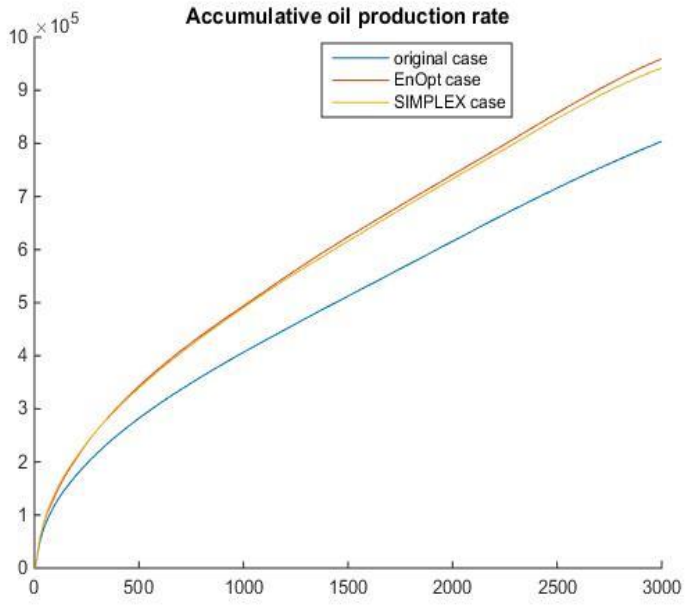

Fig 2.2.10 Comparison of Accumulative Oil Production Rate

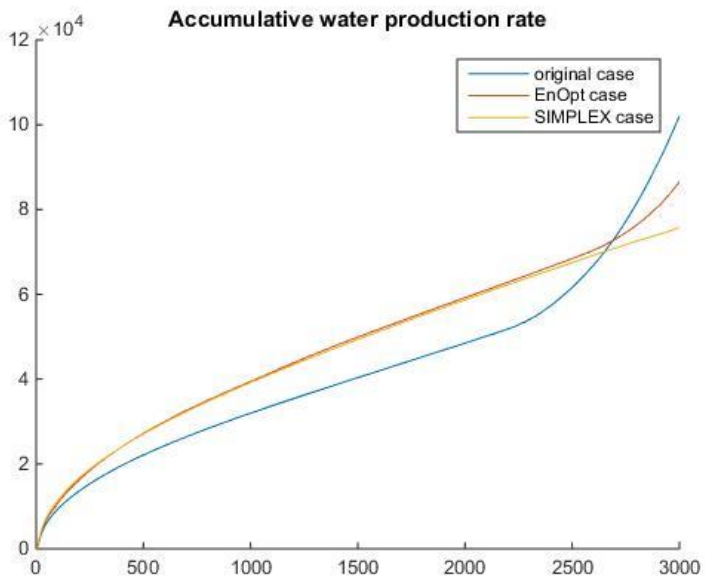

Fig 2.2.11. Comparison of Accumulative Water Production Rate

The performance of BFGS and CG with pure Simplex gradients is compared with smooth Simplex and EnOpt schemes. In Fig 2.2.12, it is observed that the CG algorithm has the best convergence rate compared to the other method; despite the algorithm is easily trapped in the local maximum. Because we only have the stochastic approximate gradient, BFGS has poor performance in both convergence rate and final optimal NPV. Compared with other optimization schemes comprehensively, the steepest ascent with the EnOpt gradient has the best performance in both convergence rate and ultimate NPV.

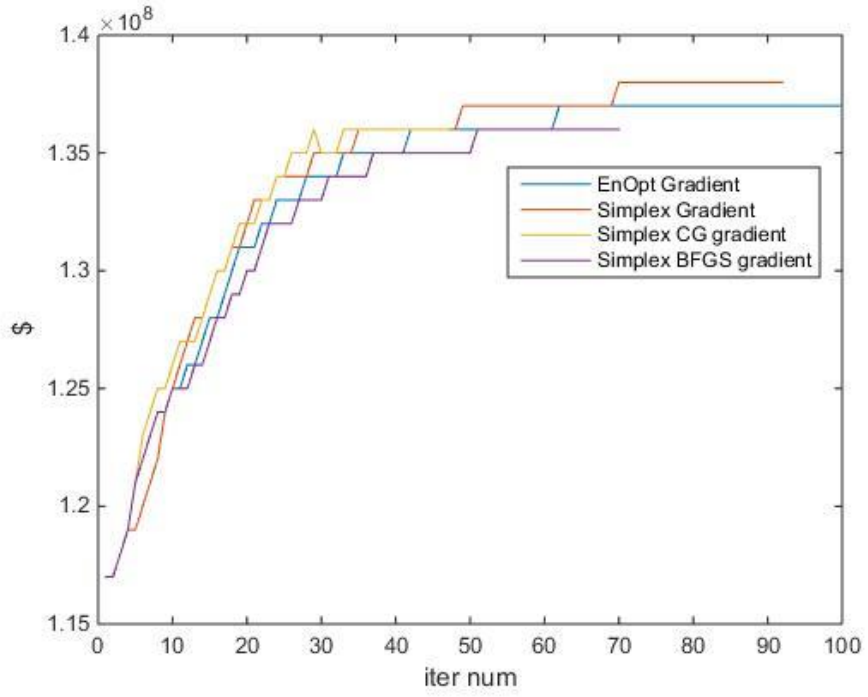

Fig 2.2.12.Comparison of Change of NPV for Different Optimization Scheme

CHAPTER 3

ROBUST OPTIMIZATION IN NFRS

Most work on production optimization (Asheim 1988; Zakirov el al.1996; Brouwer and Jansen et all. 2004; Brouwer et al. 2004; Alhuthali et al.2007; Sarmar et al.2005; Nwaozo 2006; Sarma et al.2006; Sarma et al.2008; Jansen et al.2009; Wang et al.2009) is based on a single reservoir model, which may be the mean of the ensemble of reservoir models or any single reservoir model from the ensemble. Throughout, optimization based on a single reservoir model will be referred to as nominal production optimization. Because the reservoir model may be far from the real geology, applying the optimal controls obtained from a single model to the actual field may not achieve the maximum NPV. To reduce the risk arising from the uncertainty in the geological description, van Essen et al. (2009) proposed to optimize the expectation of NPV over a set of reservoir models. This procedure is referred to as robust optimization. Van Essen et al. (2009b) compared robust optimization from more than 100 geological realizations with nominal optimization using a channelized synthetic reservoir example. They concluded that robust optimization yields not only a higher expected NPV but also a significantly smaller variance in NPV than is obtained with nominal optimization when the optimal controls obtained from nominal optimization are implemented for the set of 100 realizations.

In naturally fractured reservoirs, the spatial distribution of natural fractures is inherently uncertain. In practice, it is impossible to predict the exact location, length, and permeability of natural fractures. Robust optimization has been introduced to handle the

geological uncertainty that results from uncertain fracture networks. This uncertainty is corresponds mainly to fracture location, length, orientation and other properties (fracture permeability, aperture, and porosity) associated with fracture-fracture and fracture-matrix transmissibility. Although robust optimization is a useful method to help us handle the uncertainty introduced by natural fracture, it would take a significant amount of computational cost to get optimal well control rates for all realizations. To avoid redundant ensemble runs and to improve the efficiency of the robust optimization algorithm, the ensembles have been clustered via several proposed algorithms. By clustering ensembles properly, computational performance could improve 5-6 times compared with the standard robust optimization algorithm.

Traditional EnOpt and improved EnOpt robust optimization algorithms have been introduced. Then the k-means and hierarchical clustering have been briefly introduced as clustering algorithms to group different ensemble models into several clusters. To group ensembles reasonably, three sets of features have been selected and computed as input data for the clustering algorithms.

3.1 Robust Optimization

3.1.1 Ensemble Optimization

Robust optimization involves optimizing the expectation of the objective function with uncertain forward models. The objective function for robust optimization problems can be defined as

$$Max\{E[NPV]\},$$

$$\text{s.t. } u_{low} < u < u_{up}, \tag{15}$$

where

$$Max\{\frac{1}{N_m}\sum_{j=1}^{N_m} NPV_i\}, \tag{16}$$

the search direction is proposed as

$$d_k = C_U C_{Uf}, \tag{17}$$

where

$$Cuf = \frac{1}{N_m - 1}\sum_{j=1}^{N_m}\left(\hat{u}^j - \overline{\hat{u}}\right)\left(J\left(\hat{u}^j\right) - \overline{J}\right), \tag{18}$$

where C_U is the time dependent covariance matrix for each well, which is proposed in Appendix A.

The original robust ensemble optimization was introduced by Chen and Chen and Oliver.

A modified formulation for EnOpt for robust optimization showed as followed:

$$Cuf = \frac{1}{N_m * N_e}\sum_{i=1}^{N_p}\sum_{j=1}^{N_m}\left(\hat{u}_i{}^j - u^{j-1}\right)\left(J\left(\hat{u}_i{}^j\right) - J^{j-1}\right) \tag{19}$$

Compared with the original version of EnOpt, the modified has better convergence for well controls optimization problem in naturally fractured reservoirs.

3.2 Robust Optimization with Clustered Model

3.2.1 K-means Clustering

The K-means method is one of the simplest unsupervised learning algorithms that is used to solve the well-known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k

clusters) not are fixed a priori. The main idea is to define k centers, one for each cluster. These centers should be placed in a cunning way because of different locations lead to a different results. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the nearest center. When no point is pending, the first step is completed, and an early group age is done. At this point we need to re-calculate the k new centroids as the centers of the clusters resulting from the previous step. After these k new centroids are computed, a new binding has to be done between the same data set points and the nearest new center. A loop has been generated. As a result of this iteration we may notice that the k centers change their location step by step until no more changes are necessary. Finally, this algorithm aims at minimizing an objective function known as the squared error function given by

$$J(V) = \sum_{i=1}^{c} \sum_{j=1}^{c_i} \left(\left| \left| x_i - v_j \right| \right| \right)^2 ,$$

(20)

where, $||x_i - v_j||$ is the Euclidean distance between $x_i$and v_j, c_i is the number of data points in i^{th} cluster, and c is the number of cluster centers.

Let X =$\{x_1, x_2, x_3, \ldots\ldots x_n\}$ be the set of data points and V = $\{v_1, v_2, v_3, \ldots\ldots v_c\}$ be the set of centers.

The algorithmic steps for the k-means clustering are:

1) Randomly select c cluster centers.

2) Calculate the distance between each data point and the cluster centers.

3) Assign the data point to the cluster center whose distance from the cluster center is the minimum of all the cluster centers.

4) Recalculate the new cluster center

$$v_i = \frac{1}{c_i}\sum_{j=1}^{c_i} x_i, \tag{21}$$

where, c_i represents the number of data points in i^{th} cluster.

5) Recalculate the distance between each data point and the newly obtained cluster centers.

6) If no data point was reassigned then stop, otherwise repeat from step 3).

3.2.2 Hierarchical Clustering Method

In data mining and statistics, hierarchical clustering (also called hierarchical cluster analysis or HCA) is a method of cluster analysis that seeks to build a hierarchy of clusters.

To decide which groups should be combined (agglomerative step), or where a group should be split (divisive step), a measure of dissimilarity between sets of observations is required. The metric has been designated as a measure of distance between pairs of observation. Moreover, the linkage criterion determines the distance between sets of observation as a function of the pairwise distance between observations. In most methods of hierarchical clustering, reliable clustering is achieved by the use of an appropriate metric (a measure of distance between pairs of observations), and a linkage criterion that specifies the dissimilarity of sets as a function of the pairwise distances of the observations in the sets.

In this work, the Euclidean distance (two normal) has been chosen as the clustering metrics to measure the distance between observations.

In this work, the average linkage clustering criterion has been selected. Given two observation $a \epsilon A$ and $b \epsilon B$ is the sets which cardinality $|A|$ and $|B|$ respectively, the criterion is:

$$\frac{1}{|A||B|} \sum_{a \epsilon A} \sum_{b \epsilon B} d(a, b) \tag{22}$$

3.2.3 Selection of Features and Work Flow for Clustering Optimization

The most intuitive way to cluster the fracture models is to group different fracture models via their resulting NPV. The workflow for NPV clustering is:

1) Simulate whole ensembles with initial wells controls.

2) Rank the NPV for each ensemble.

3) Select the ensembles correspondent with $P_{10}, P_{20}, P_{30} \ldots P_{90}$ in rank NPV.

4) Perform robust optimization with selected models.

The workflow to cluster fracture models based on model parameters is:

1) Select and compute model parameters based experience.

2) Cluster the ensemble via k-means or the hierarchical method into n clusters.

3) Select one model from each group.

4) Run robust optimization with selected models as the ensemble.

To describe the features of each fracture model, three model parameters, such as total conductivity, total fracture storability, and weighted fracture orientation have been defined. The formulations to compute model parameters for each fracture model are showed as follows:

$$S_{total} = \sum_{i=0}^{n_f} \phi_i h_i l_i A_i, \tag{23}$$

$$Con_{total} = \sum_{i=0}^{n_f} k_i A_i l_i, \tag{24}$$

and

$$O_{weighted} = \sum_{i=0}^{n_f} \varphi_i k_i A_i l_i. \quad (25)$$

Clustering different ensembles via response, such as accumulative oil, water production rate and the breakthrough time corresponding to each well segment and NPV with respect to each model is reliable ways to group model. This strategy is more reliable when the geological uncertainty such as uncertainty in permeability and porosity field has been introduced in each ensemble model.

1) Simulate whole ensemble and record response for each model.

2) Clustering whole ensemble via the hierarchical clustering and responses of each model.

3) Select one model from each cluster.

4) Performance robust optimization with selected ensembles.

3.3 Computational Results

In this section, two examples have been selected to illustrate the performance of our clustered robust optimization framework. In Example 1, smart wells control are optimized for the NFR with homogeneous matrix properties. The results of clustered robust optimization are compared with standard robust optimization without clustering. In Example 2, geological uncertainty is introduced into the reservoir model and the smart well controls are optimized for cases with fracture distribution uncertainty and with geological uncertainty. In all the examples, the objective is to maximize the NPV by optimizing the BHP for the production well and the injection rates for the injection well.

The basic reservoir model is a two-dimension, two-phase (oil-water) homogenous NFR. The reservoir size is 2500ft by 2500ft and a simulation grid of dimension 50x50x1

is used. The permeability and porosity fields are homogenous with values 10md and 0.2 respectively. Two smart horizontal wells are drilled parallel to each other. One is a producer and the other is an injector. The heel and toe coordinates of the production well are (125, 250) and (125, 2250) respectively. The location of the injection well is defined by the coordinates (2325, 250) and (2325, 2250) corresponding to the start and end points. Each well is divided into 8 equal segments and each segment is independently controlled by an ICV.

Upper and lower bounds have been applied to constrain the BHP for producers and injection rates for injectors. The maximum BHP is 4000 psi, which is equal to the initial reservoir pressure. The minimum BHP is set to 1000 psi for the producer. The maximum and minimum injection rates are set as 500 and 0 STB per day respectively. The initial production pressure and injection rate have been set at 2000 psi and 100 STB/d. In the robust optimization case, the oil price has been set as $80/STB. The injection water and disposal water price are $3/STB and $5/STB.

	Mean Value	Standard Deviation	Units
Number of Fracture	20	3	
Fracture Length	800	50	ft
Fracture Orientation	0	0.393	rad
Fracture Permeability	2000	200	md
Fracture Aperture	0.15	0.02	ft
Fracture Porosity	0.95	0	
Number of Model	200	0	

Tab 3.1 The Basic Parameters to Generate Fracture Fields

3.3.1 Robust Optimization with Homogenous NFRs

In order to compare the performance of the robust optimization algorithm with and without different clustering techniques, four numerical scenarios have been designed. In the first scenario, 100 natural fracture models have been generated to represent the underground uncertainty. The standard robust optimization using full ensembles is performed. In the second scenario, the full ensemble is run once and the $P_{10}, P_{20}, P_{30} \ldots P_{90}$ are selected based on rank their of NPV. The robust optimization algorithm is performed using the selected models. In the third scenario, the full ensembles are clustered based on the model parameters defined in section 3.2.3, such as total fracture conductive, total fracture storability and weighted fracture direction. The clustering results showed in Figure B.1. In Appendix B Fig B.1, the numbers on the x-axis represent the label of the model. The length on the y-axis represents the normalized distance between two different models or clusters. By setting up the normalized threshold distance as 0.2, we could group whole ensembles into 15 clusters.

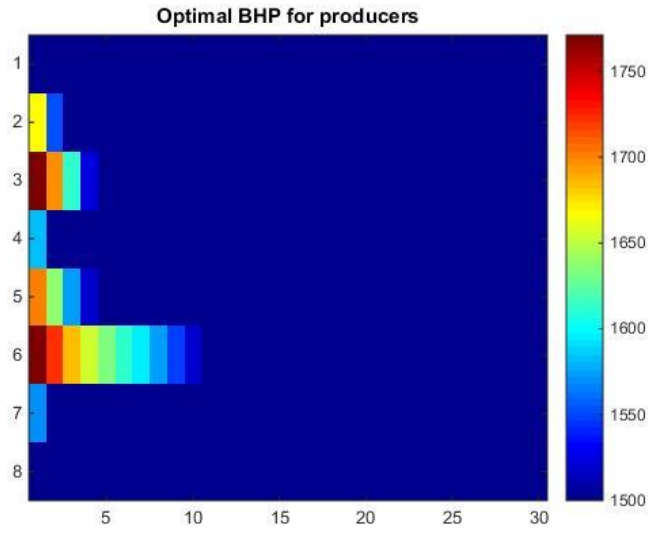

Fig3.3.1 The Optimal BHP Controls for Production Well Segment (Full Ensemble)

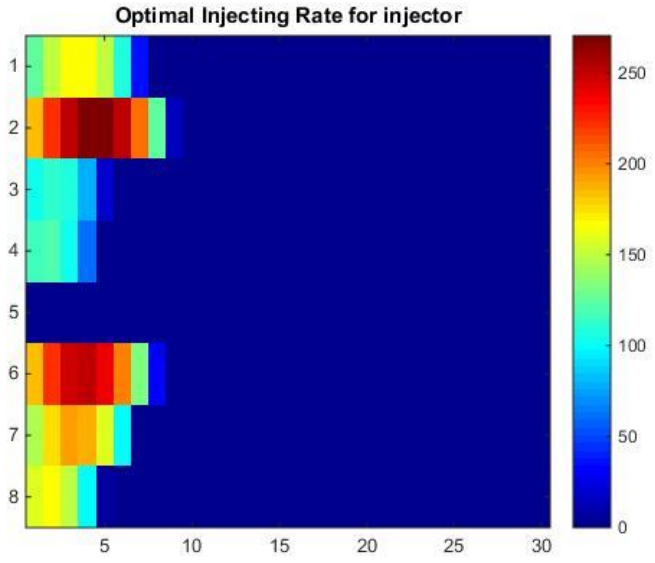

Fig 3.3.2 The Optimal Injecting Rate Control for Injection Well Segment (Full Ensemble)

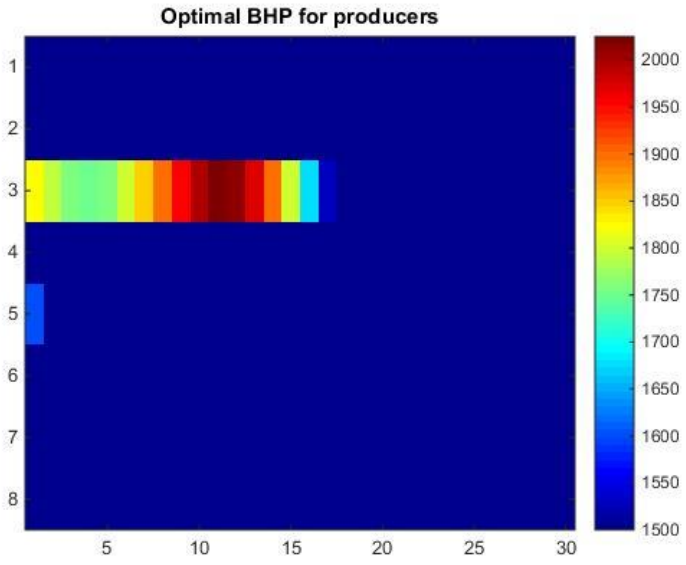

Fig3.3.3 The Optimal BHP Controls for Production Well Segment (Selected NPV)

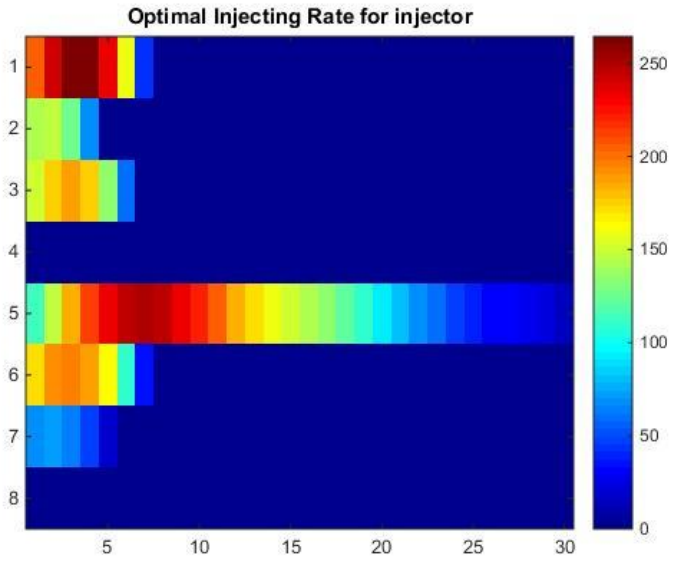

Fig 3.3.4 The Optimal Injecting Rate Control for Injection Well Segment (Selected NPV)

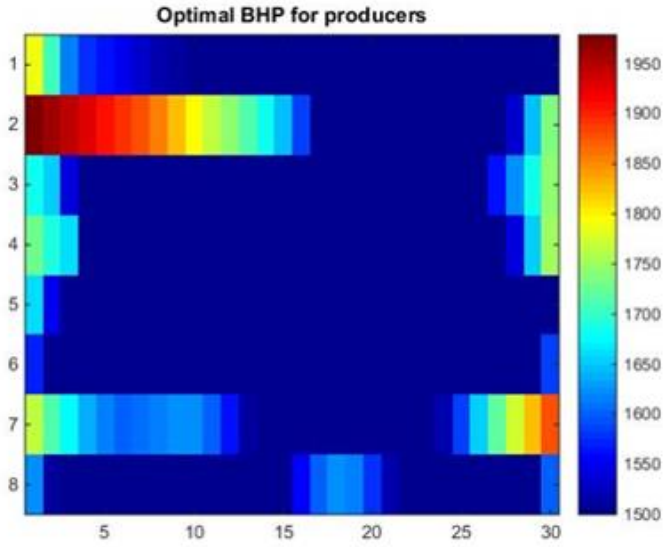

Fig3.3.5 The Optimal BHP Controls for Production Well Segment (Model Clustering)

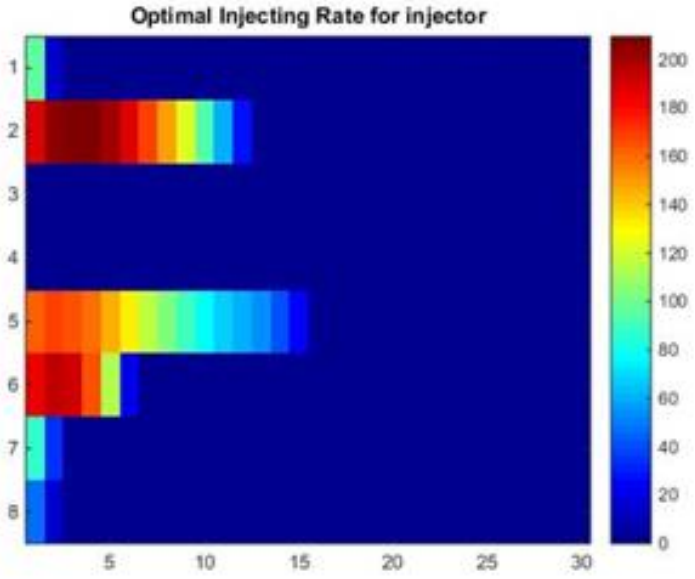

Fig 3.3.6 The Optimal Injecting Rate Control for Injection Segment (Model Clustering)

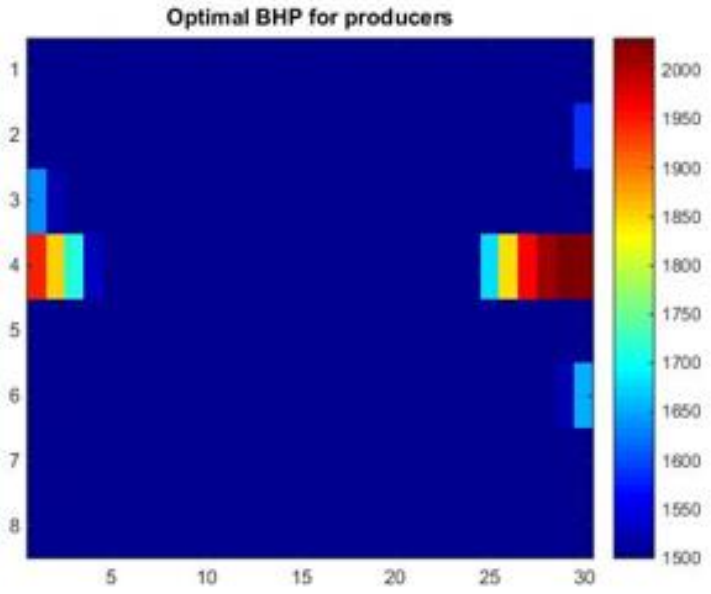

Fig3.3.7 The Optimal BHP Controls for Production Well Segment (Response Clustering)

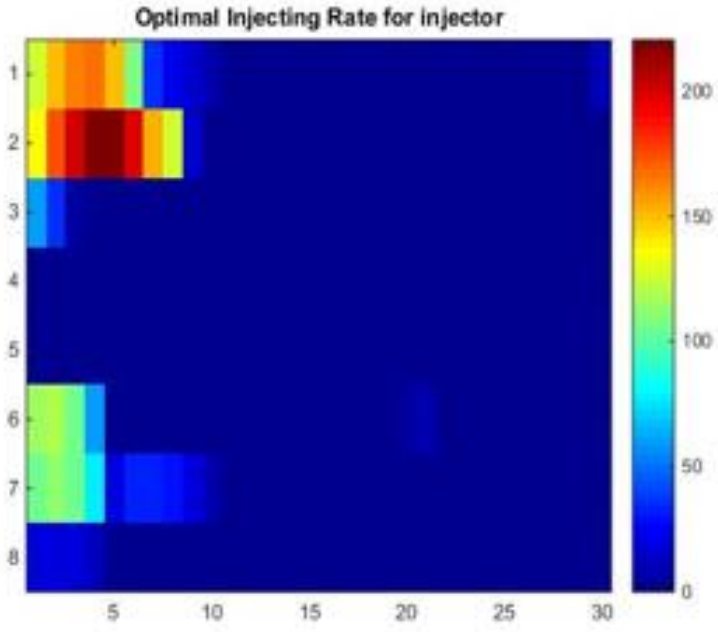

Fig 3.3.8 The Optimal Injecting Control for Injection Segment (Response Clustering)

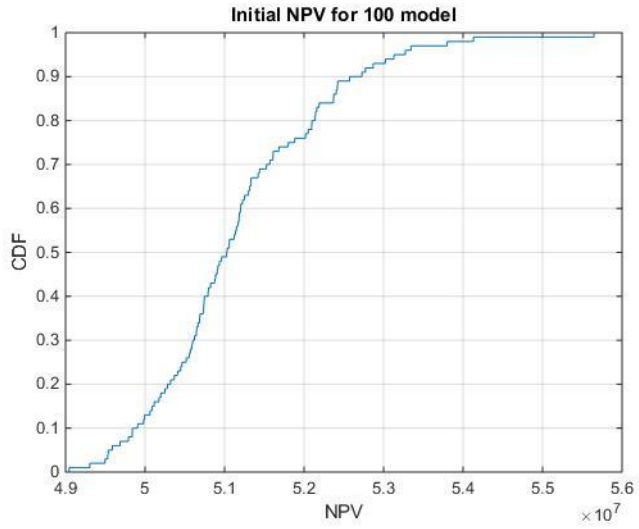

Fig 3.3.9 Initial NPV of Full Ensemble

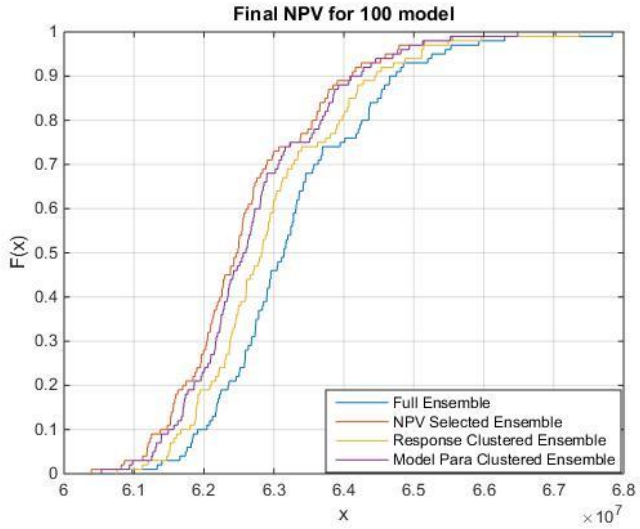

Fig.3.3.10 Final NPV of Full Ensemble

Cases	Std	Mean	Max NPV	Min NPV	Num of Sim
Full Ensemble	1148900	63281160	67838000	60969200	2500
Selected Ensemble	1125700	62622471	67030700	60388000	340
Clustered by Model	1064200	62731600	66486700	60532100	430
Clustered by Response	1081300	62978000	67366400	60753100	400

Tab 3.2 Comparison of Different Clustering Schems

In Figs 3.3.1, 3.3.3 and 3.3.5, it can be observed that the optimal well controls always lead the operator to restrict BHP and water injection at the beginning of production. However, when comparing with the initial reservoir pressure (4000psi), there are only minor restrictions on BHP at the beginning of production. When it comes to injectors, we could observe that the injection wells always tend to inject water at the beginning of the production period. Injecting water at the beginning of production not only improves the efficiency of the water flood but also prevents early breakthrough of water. By comparing different well controls generated by different clustering robust optimization algorithms, it easy to recognize that there is some similarity between the resulting optimal well controls. For example, all optimal controls tend to restrict water injection in segments 3 and 4. Moreover, all optimal well controls lead us to inject a significant amount of water at well segments 2 and 5. By comparing the final CDF of NPV for the whole ensembles, we could compare the performance of different clustering robust optimization methods. From the Fig 3.3.10, it easy to recognize that by using the full ensemble robust optimization, a better NPV is obtained than by any other clustering optimization methods. However, the full ensemble robust optimization could be computationally inefficient takes a number of simulations to reach the optimal well controls. When it comes to the proposed clustering algorithms, the response clustered robust optimization has the highest expected NPV. Additionally, the standard deviation of NPV for the response clustered optimal well controls is also smaller than for other methods, except the model parameters clustered method. From the Table3.2, it could be observed that the standard optimization algorithm takes 2000 of simulation runs to reach the optimal well controls; however, the response clustering robust optimization algorithm only takes 400 simulations to find the optimal

well controls. By using the response clustering method, we could improve the efficiency of robust optimization 5-6 times and without sacrificing expected NPV for whole ensembles. Generally speaking, the response clustering method has the best performance in this case.

3.3.2 Robust Optimization with Geological Uncertainty

In this example, besides the uncertainty introduced by the random fracture distribution, the geological uncertainty in permeability and porosity fields is considerate. Because of geological uncertainty, it would be unreliable if we only group the fracture models by the fracture model parameters. In this section, only three different robust optimization algorithms have been applied, which are the standard robust optimization, the NPV selected optimization and the responses clustered optimization.

In comparison with the previous example, we leave all parameters unchanged except that the matrix permeability and porosity fields. Moreover, because there is more uncertainty that is introduced into the system, the number of the ensemble is increased to 200 instead of 100. The permeability fields are assumed to obey log-normal distribution whereas the porosity fields obey normal distribution. The 200 realizations of matrix permeability and porosity fields are generated using a spatial covariance matrix and Gaussian random noise. The formulation of the covariance matrix is illustrated in Appendix A. The mean values of log permeability and porosity are set as 2.3 md and 0.2 respectively, and the standard deviations for log-permeability fields is 0.3. The standard deviation for matrix porosity is 0.03. The maximum and minimum correlation lengths for the covariance matrix are 1500ft and 300ft. The cross-correlation coefficient between permeability and

porosity is 0.5. The orientation angle for the direction of maximum correlation is $\frac{\pi}{4}$. To demonstrate the permeability and porosity distribution of the ensemble, 3 realizations have been selected and appear in Fig 3.3.11-3.3.13. From these figures, it easy to recognize that there are some high permeability zone that are distributed randomly along the $\frac{\pi}{4}$ direction.

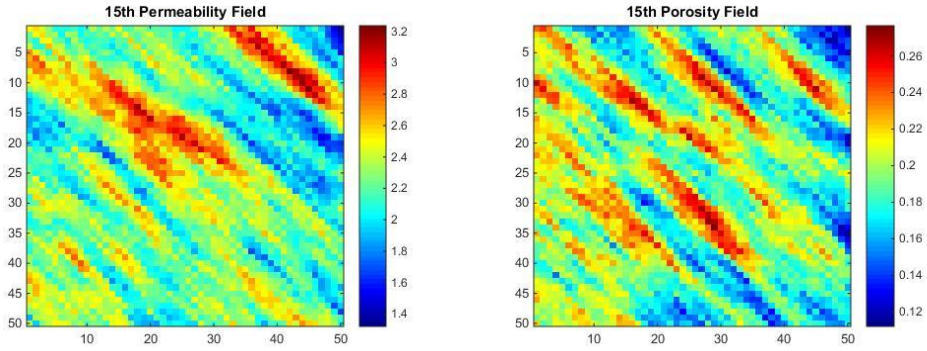

Fig 3.3.11 Permeability and Porosity Distribution for 15th Model

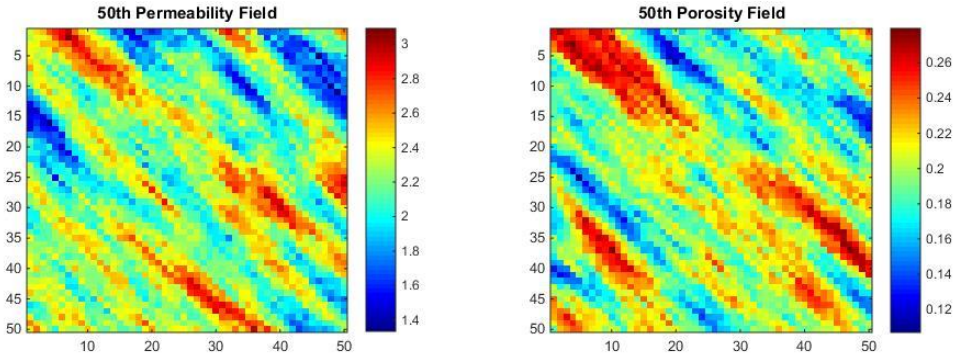

Fig 3.3.12 Permeability and Porosity Distribution for 50th Model

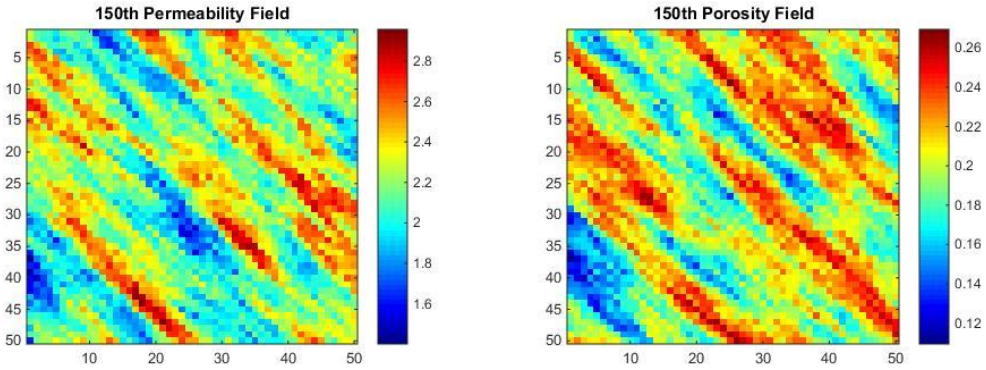

Fig 3.3.12 Permeability and Porosity Distribution for 150th Model

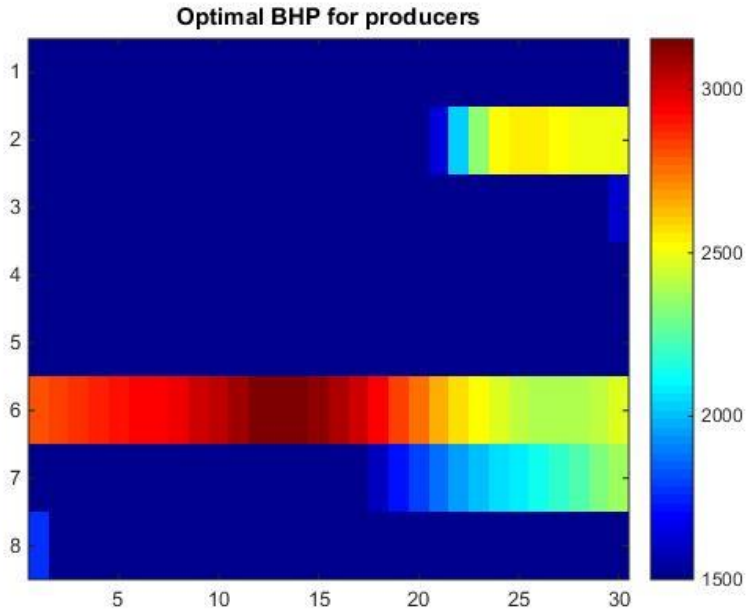

Fig 3.3.13 The Optimal BHP Controls for Production Well Segment (Full Ensemble)

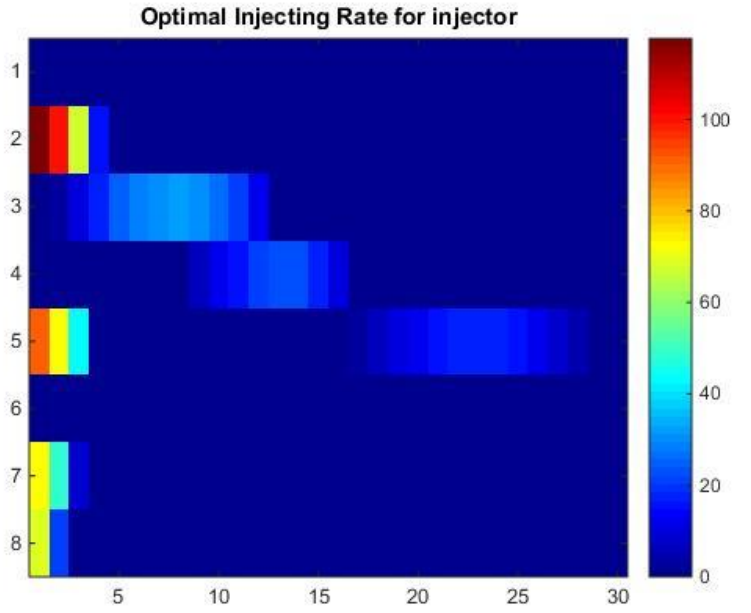

Fig3.3.14 The Optimal Injecting Controls for Injection Well Segment (Full Ensemble)

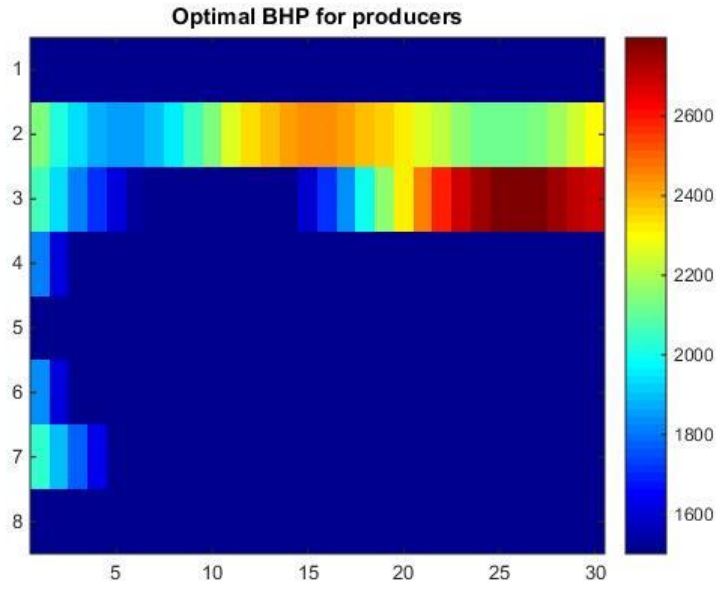

Fig 3.3.15 The Optimal BHP Controls for Production Well Segment (NPV Selected)

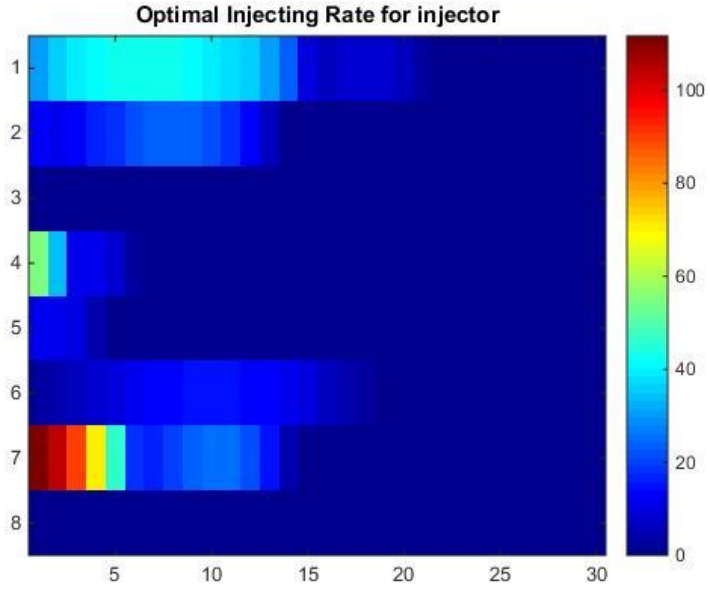

Fig 3.3.16 The Optimal Injecting Controls for Injection Well Segment (NPV Selected)

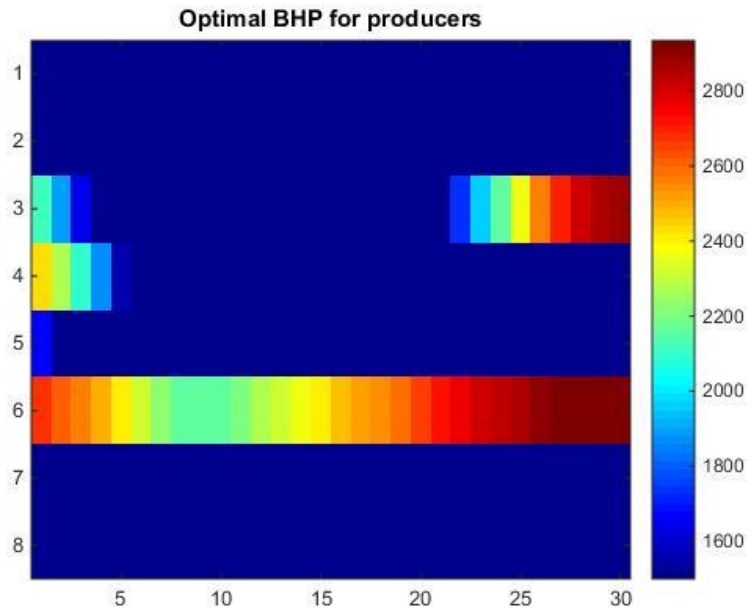

Fig 3.3.17 The Optimal BHP Controls for Production Segment (Response Clustering)

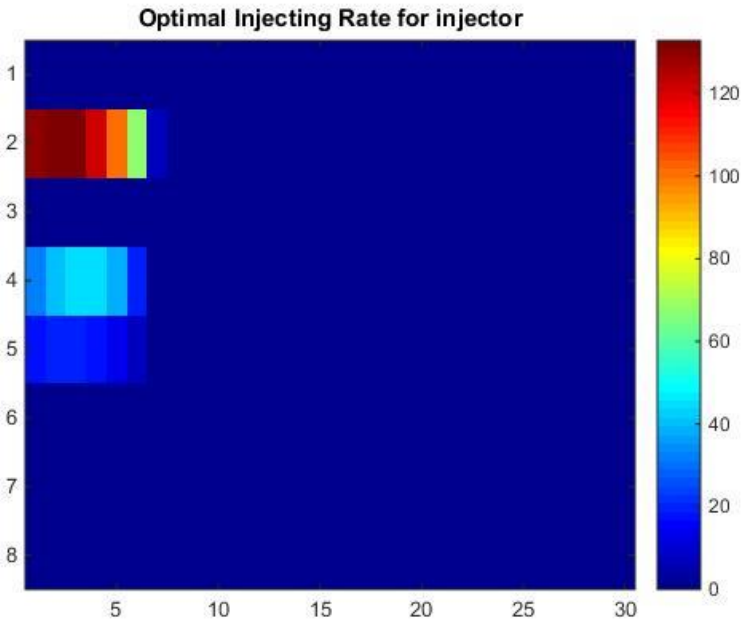

Fig 3.3.18 Optimal Injecting Controls for Injection Well Segment (Response Clustering)

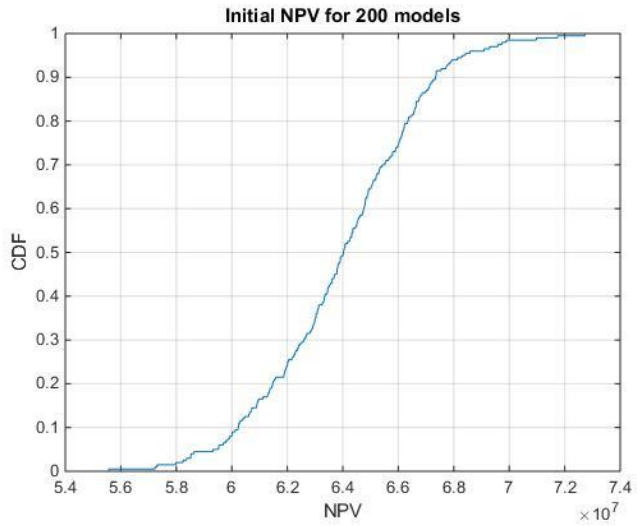

Fig 3.3.19 Initial NPV of Full Ensemble with Geological Uncertainty

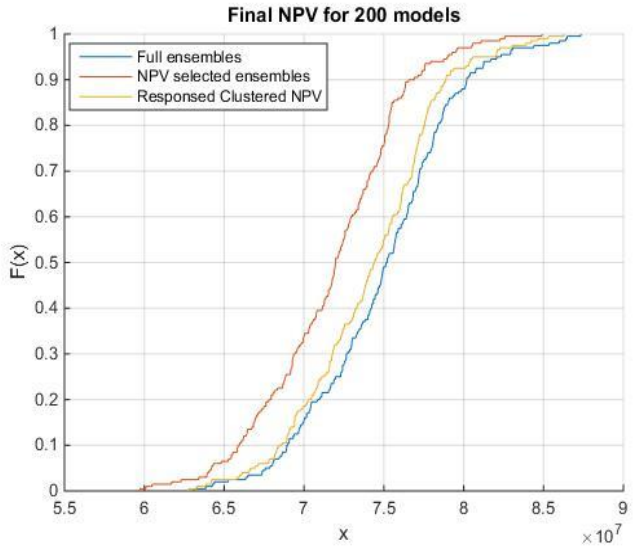

Fig.3.3.20 Final NPV of Full Ensemble with Geological Uncertainty

By comparing the optimal well controls generated by the full ensemble and the response clustered ensemble, several similarities between the optimal well controls are observed. In Figs 3.3.13 and 3.3.17, the 6^{th} production segment continuous restricts production by controlling the BHP during entire production life. As for injectors, it is observed that both schemes tend to inject water at the beginning of the production. Injecting

water at the beginning of production can not only improve the efficiency of flood but also prevent the breakthrough of water front. Moreover, both control schemes tend to inject water at segments 2 and 5. In Fig 3.3.20, the CDF of NPV using the various clustering methods show that by using the standard ensemble robust optimization, we could get better NPV than the NPV selected optimization algorithm. However, the full ensemble robust optimization could be computationally inefficient and it takes an amount of simulations to reach the optimal well controls. The standard robust optimization with 200 ensembles takes 5400 simulations to reach the optimal NPV and the NPV selected optimization only needs 850 simulations to converge. Although NPV selected optimization has a good performance on efficiency, the final expected NPV for full ensemble is much smaller than other methods, which means the optimal control for selected model might not be the optimal control for full ensemble. As to the response clustering optimization algorithm, it only takes 650 simulation run to reach optimal expected NPV. Hence from Fig 3.3.20, it is easily concluded that the response clustering optimization algorithm could improve the efficiency of the robust optimization significantly without sacrificing the final expected NPV for full ensemble.

CHAPTER 4

DATA ASSIMILATION AND HISTORY MATCHING

Continuum fracture models, such as the Dual Porosity (DP) and the Dual Permeability and Dual Porosity (DP-DK) models have the advantage over DFM in that the fracture location and properties can be altered by changing the fracture cell permeability, shape factor, and pore volume. The final fracture permeability, shape factor, and pore volume fields represent a probability that fractures will appear. For example, a certain region with high permeability and pore volume might indicate there is a high probability for this region to have natural fractures.

In NFRs, the observed data is generated by applying observation error on true production data, and the true production data is generated by a homogenous field with randomly generated EDFM model. The prior ensemble models are generated by applying Gaussian noise on the spatial covariance matrix. The details of the covariance matrix are illustrated in Appendix A.

4.1 Models and Methods

4.1.1 Dual Permeability Dual Porosity Model

The DP model proposed by Warren and Root (1963) is a widely used method to simulate fluid flow in the naturally fractured reservoir. Although DP models are limited in their ability to represent complex fracture geometry, however, when only we have a probability distribution of a naturally fractured field, the DP models would have a better

representation of the fracture field via changing the value of shape factor and fracture permeability. The connection list for DP model is showed in Fig 4.1.1. From Fig 4.1.1, we could observe that in the DP model, there is no direct connection between matrices and each matrix cell had been connected with each other by a set of fracture cells. The dual porosity dual permeability (DP-DK) is an enhancement to the standard DP model. In this model, communication between the matrixes (the inter-granular void space which is also referred to as the primary porosity) is not assumed to be negligible. The inter-block communication for the DP-DK model shown in Fig 4.1.2.

Oil in the fracture is conserved:

$$\Delta T_{of}^{x}\left(\Delta p_{of}^{n+1} - r_{of}^{x}\Delta D\right) + \Delta T_{omf}^{x}\left(p_{of}^{n+1} - p_{om}^{n+1}\right) - \frac{V_b}{\Delta t}\left[\left(\frac{\phi S_o}{B_o}\right)_f^{n+1} - \left(\frac{\phi S_o}{B_o}\right)_f^{n}\right] = 0 \qquad (26)$$

Water in the fracture is conserved:

$$\Delta T_{wf}^{x}\left(\Delta p_{wf}^{n+1} - r_{wf}^{x}\Delta D\right) + \Delta T_{wmf}^{x}\left(p_{wf}^{n+1} - p_{wm}^{n+1}\right) - \frac{V_b}{\Delta t}\left[\left(\frac{\phi S_w}{B_w}\right)_f^{n+1} - \left(\frac{\phi S_w}{B_f}\right)_m^{n}\right] = 0 \qquad (27)$$

$$T_{of}^{x}{}_{i+\frac{1}{2}} = \left[\frac{\Delta Y \Delta Z}{\Delta X} k_e\right]_{i+\frac{1}{2}} \cdot \left[\frac{k_{ro}}{\mu_o B_o}\right]_{f_k}^{x} \qquad (28)$$

$$k_e = k_f \cdot \frac{V_{fp}}{V_b} \qquad (29)$$

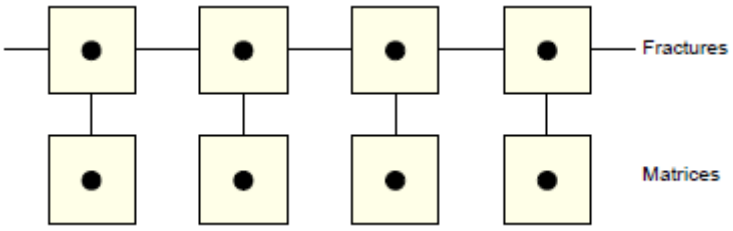

Fig 4.1.1 Schematic Diagram of Connectivity for the Standard DP Model

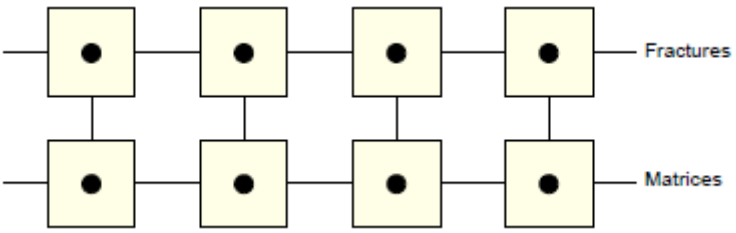

Fig 4.1.2 Schematic Diagram of Connectivity for the Standard DPDK Model

4.1.2 Ensemble Smoother with Multiple Data Assimilation

Ensemble methods use an ensemble of model realizations to represent the uncertainty in the model state and parameters. However, unlike oceanic and atmospheric models that present chaotic and unstable dynamics (Evensen, 2007, Chapter 6) reservoir simulation models are typically a stable function of the rock property field. If we make the assumption that there is no model error for the simulator, we need only estimate model parameters when applying the ensemble smoother (ES). The formulation of the ES method showed is as:

$$\mathrm{m}_j^a = \mathrm{m}_j^f + C_{MD}^f\left(C_{DD}^f + C_D\right)^{-1}\left(d_{uc,j}^f - d_j^f\right), \tag{30}$$

for $j = 1,2, \ldots, N_e$, where N_e denoting the number of ensemble members; C_{MD}^f is the cross-covariance matrix between the prior vector of model parameters, m_j^f ,and the vector of realization data d^f, C_{DD}^f is the $N_d \times N_d$ auto-covariance matrix of predicted data; and C_D is variance of predicted data, which typically is a diagonal matrix.

In Emerick and Reynolds (2012), it was shown that the single and multiple data assimilations for the linear Gaussian case. As to our reservoir simulation, which has high nonlinearity due to relative permeability capillary pressure etc., this equivalence does not hold. Inspired by Gaussian-Newton iteration, multiple data assimilation (MDA) can be

viewed as an iterative ensemble smoother; we perform multiple smaller corrections in the ensemble to overcome nonlinearity of the problem. The essential steps for ES-MDA is shown as:

1. Choose the number of data assimilation N_a, and the coefficient α_i for i=1… N_a.

2. For i=1 to N_a

(a) Run the ensemble from time zero.

(b) For each ensemble member, perturb the observation vector using $\mathrm{d}_{uc} = \mathrm{d}_{ob} + \sqrt{\alpha_i}\, {C_D}^{1/2} z_d$, where $z_d \sim N(0, I_{N_d})$.

(c) Update the ensemble using Eq. (13) with C_D replaced by $\alpha_i C_D$.

The more details of the ES-MDA algorithm are well summarized in Emerick and Reynolds (2012).

4.2 Computational Results

4.2.1 Example 1.History Matching for NFRs

In this example, the history matching work flow for NFRs has been introduced. The size of the model and basic parameters are set same as the examples in Chapter 4. The sketch of fracture distribution for the true model is shown in Fig 4.2.1. The other parameters of the true fracture model are shown in Tab 4.2.2. Then the DPDK model has been used as the forward model to assimilate the production data generated by the true model.

Setting up initial fracture fields in DPDK is also a critical step in the data assimilation workflow. The quality of prior fracture fields can deeply affect the result of the history matching. To set up prior model properly, the mean value of log fracture

permeability has been set as 7, mean value of fracture cell shape factor is 2665 and mean value of fracture pore is 20. The maximum and minimum correlation length for covariance matrix are 1000ft and 500ft. The cross-correlation coefficient between permeability and porosity is 0.5. The orientation angle for the direction of maximum correlation is 0. Figs 4.2.3 and 4.2.4 are random selected prior fields before assimilating the observed data. We could observe there is large variety in oil water production rates and water breakthrough times, which means there large amount of contained in prior models. As nonlinearity of this problem, the ensemble size is set as 200 and the assimilation number is 7. As we known, the harmonic summation of inflation rates in each assimilation should equal to 1; the inflation factors for each assimilation shown in Table.4.2.1.

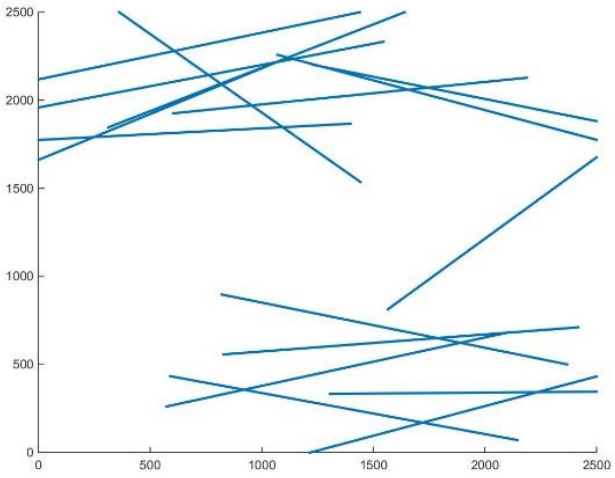

Fig 4. 2.1 Fracture Distribution of True Model

Number of Assimilation	1	2	3	4	5	6	7
Inflation Rate	3000	1000	500	200	6.862	3	2

Fig 4.2.1 The Inflation Rate for Each Assimilation

	Mean Value	Standard Deviation	Units
Number of Fracture	15	3	
Fracture Length	1200	50	ft
Fracture Orientation	0	0.392699075	rad
Fracture Permeability	2000	200	md
Fracture Aperture	0.1	0.02	ft
Fracture Porosity	0.95	0.01	
Number of Model	200	0	

Fig 4.2.2 Basic Parameters to Generate True Fracture Field

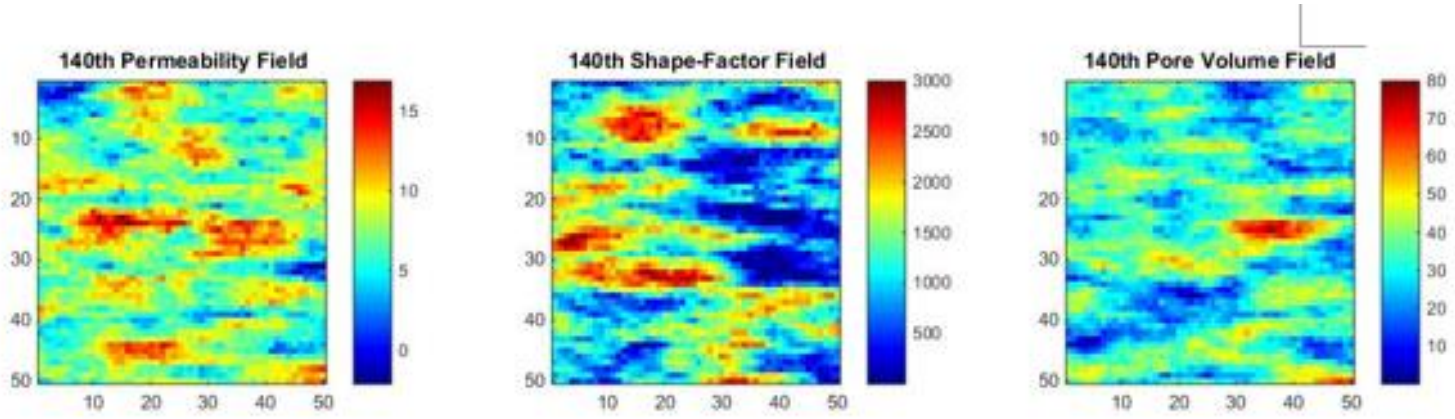

Fig 4.2.3 The 140th Initial DPDK Field

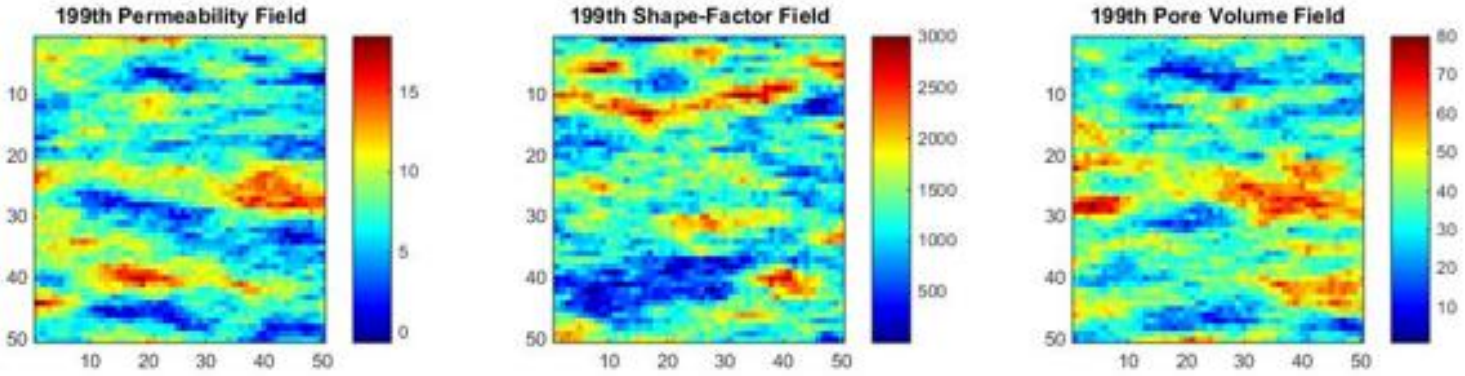

Fig 4.2.4 The 199th Initial DPDK Field

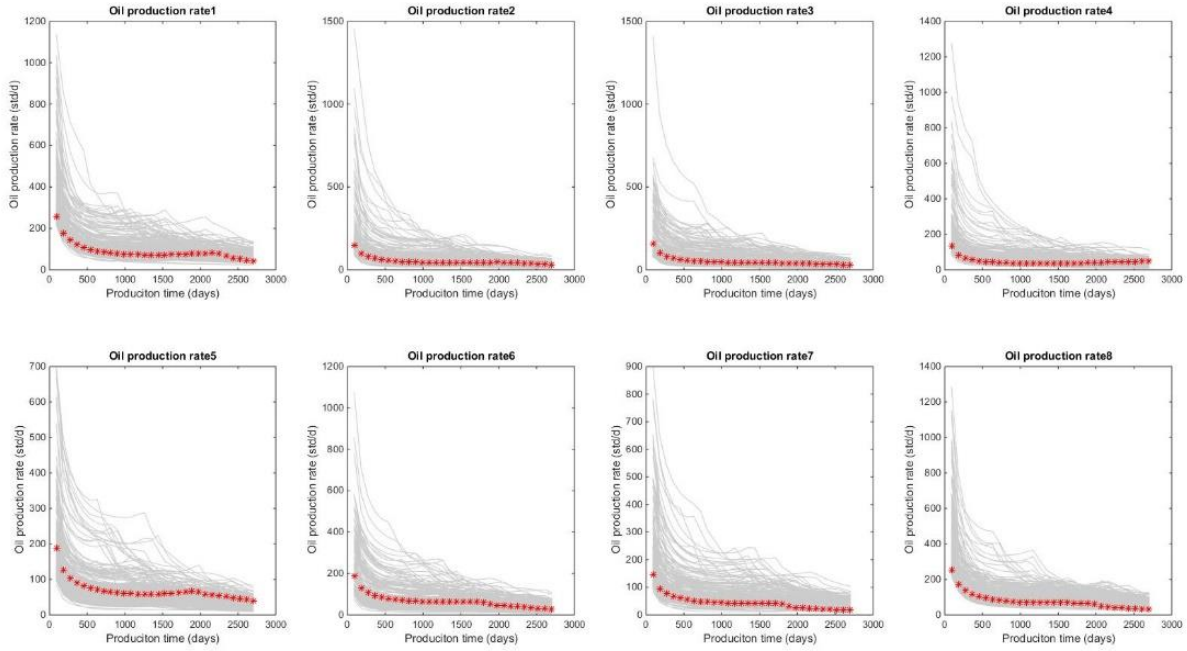

Fig 4.2.5 The Oil Production Rate of 200 Ensemble before Assimilating Data

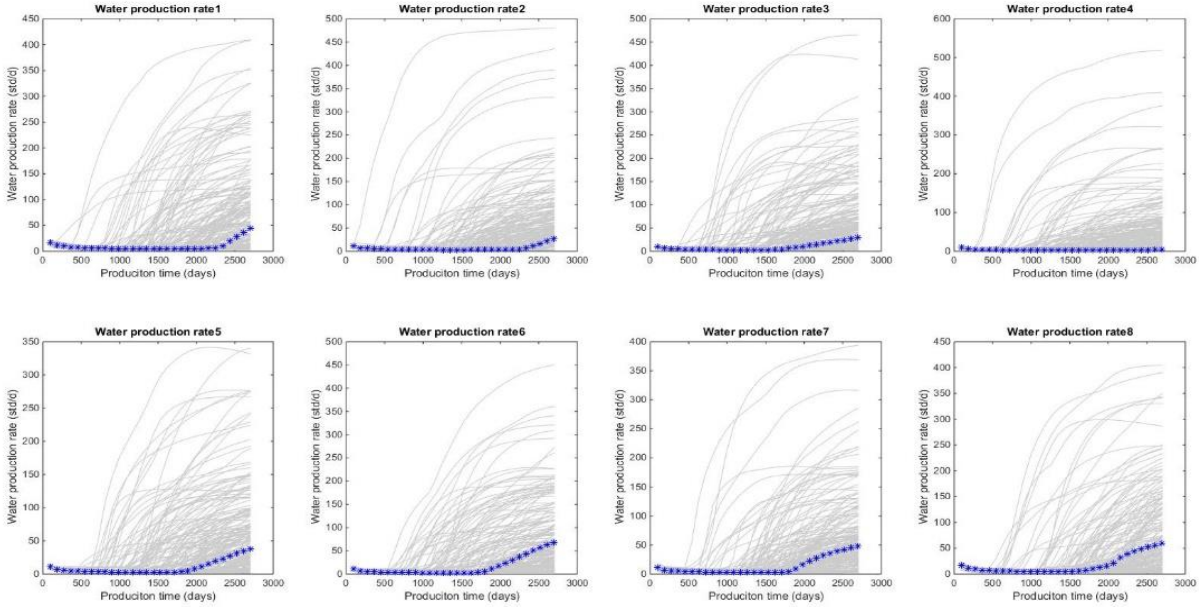

Fig 4.2.6 The Water Production Rate of 200 Ensemble before Assimilating Data

In Figs 4.2.9 and 4.2.10, we could observe that both oil and water production rates for each well segment matched with the observed data. There is little bit mismatch for water breakthrough time for well segments 3, which is caused by inconsistency between DFM and continuum fracture model. The constrained DPDK fields shown in Figs 4.2.7 and 4.2.8. By comparing constrained DPDK fields (Fig 4.2.7 and 4.2.8) with the true fracture field (Fig 4.2.1), there are some easily observed agreements between these fields. For example, the upper part and lower left part of permeability field and pore volume field have distinct large value than other parts of those fields. In Fig 4.2.1, we also observe that the upper part and the lower left part of the field have larger fracture intensity than other part of the field .Thus this example illustrates that it would be a reasonable way to use data generated by DFN to constrain the DPDK model to reduce the uncertainty of the NFRs system.

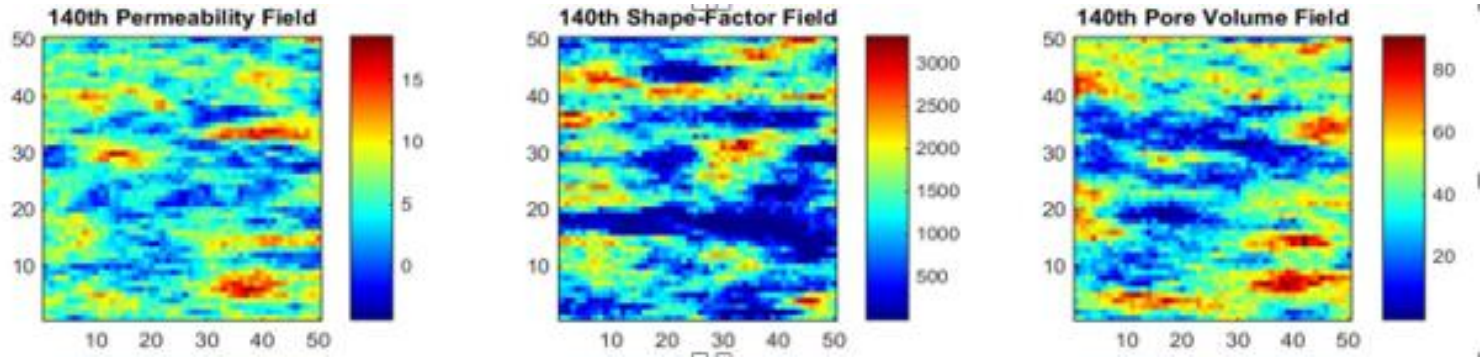

Fig 4.2.7 The 140th DPDK Field after Assimilating Data

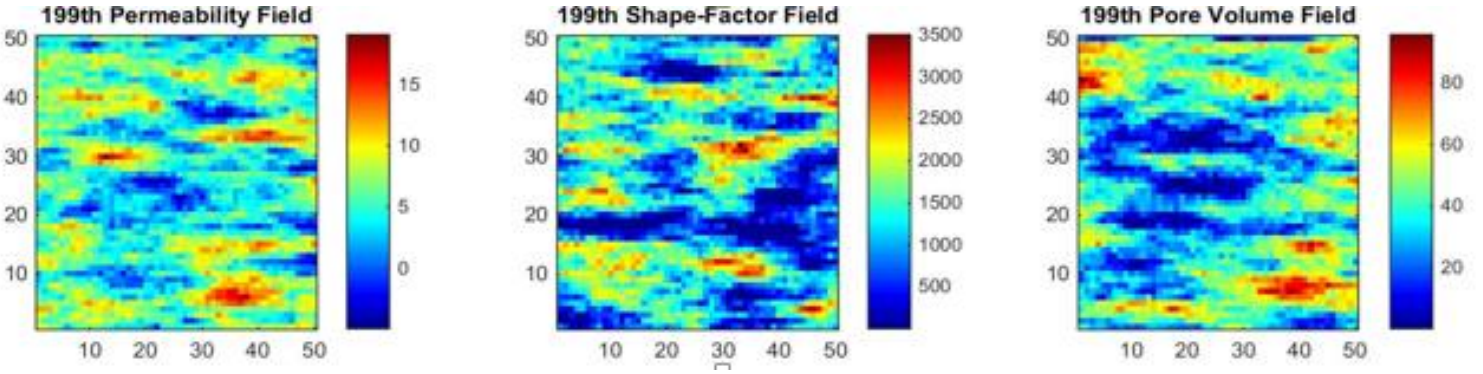

Fig 4.2.8 The 199th DPDK Field after Assimilating Data

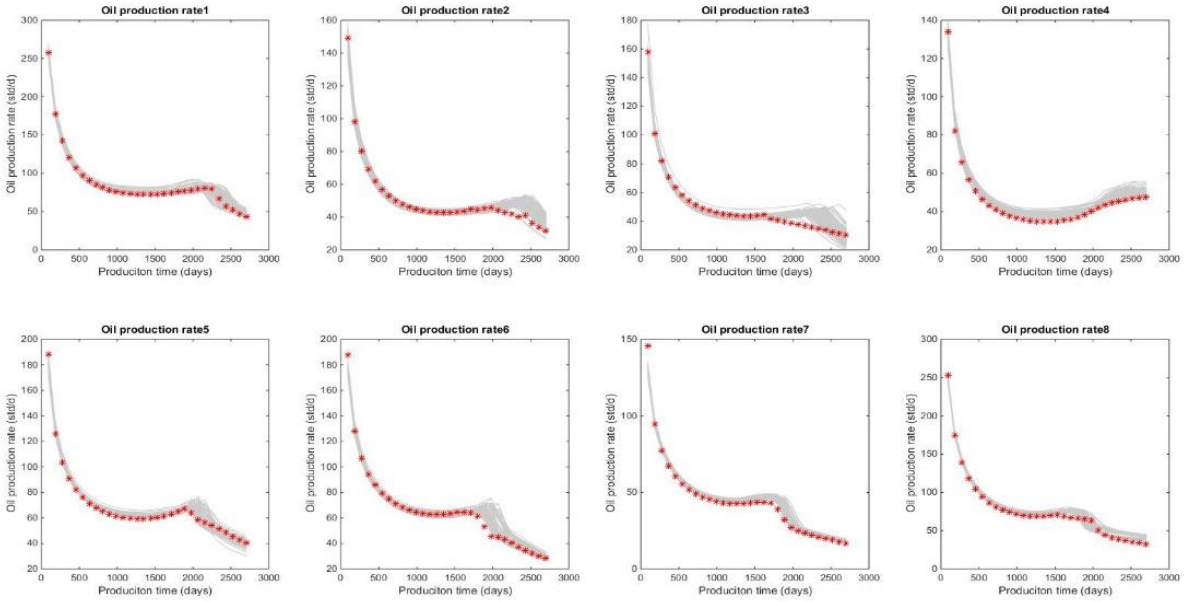

Fig 4.2.9 The Oil Production Rate of 200 Ensemble after Assimilating Data

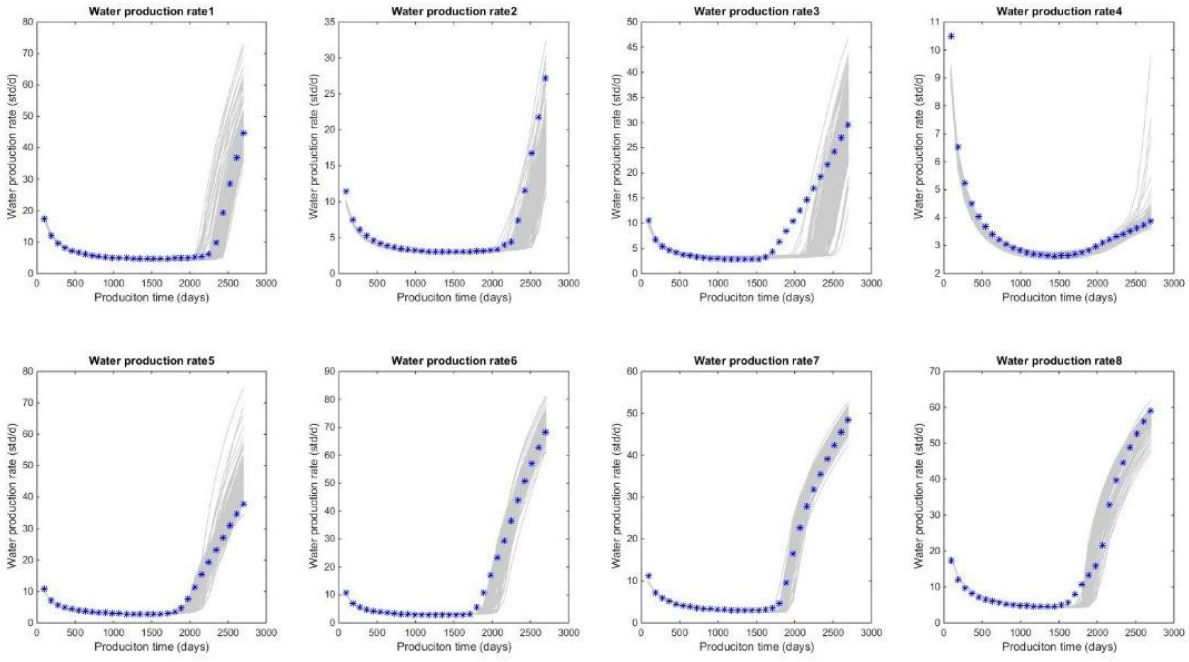

Fig 4.2.10 The Water Production Rate of 200 Ensemble after Assimilating Data

CHAPTER 5

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary and Conclusions

Production optimization for naturally fractured reservoirs is difficult because the randomly distributed natural fractures introduce nonlinearity and uncertainty into the optimization system. Optimal well controls should increase the oil production rate and meanwhile decrease the water production rate.

In Chapter 2, the basic EDFM was used to simulate fluid flow in natural fractures and the objective function for optimization algorithm has been defined. The steepest ascent with scaled variables has also been applied to maximize the objective function. It is observed that developing NFRs with optimal production rate could increase oil production rate and meanwhile restrict water production rate. The performance of the Simplex gradient with different optimization schemes has been compared with the EnOpt gradient with the steepest ascent optimization scheme. The results showed that the CG algorithm has the best convergent rate compared with other method for the cases presented but would be easier to be trapped into local maximum.

In Chapter 3, a clustering-based robust optimization procedure is developed. The optimization procedure consists of robust optimization and clustering of different models in two steps. Three kinds of different features, such as selected NPV, model parameters clustering and model responses clustering, have been chosen to cluster the numerical models. The k-means method and the hierarchical method are also introduced as clustering

methods to group the models. The computational results show that although the robust optimization has the best performance in expected NPV aspect than other clustering methods, it would be computationally expensive. The response clustering optimization not only has a good representation of ensemble uncertainty but also can improve the efficiency of robust optimization significantly.

Chapter 4 demonstrates a workflow to history match NFRs. In this workflow, the DPDK model has been used as the forward model to match the observation data, which is generated by fracture fields with EDFM. From the computational results, we could observe that there are good matches for oil production rate and water production rate. Moreover, the regions of large fracture intensity in the true model, tend to have large pore volume and fracture permeability in constrained DPDK field. The result shows there is a good agreement in fracture distribution in both EDFM and DPDK models.

5.2 Future Work

The response clustering method is, we believe, quite promising though several aspects require additional investigation. These include the following:

1. Although the total fracture conductivity, storability and weight fracture orientation have been selected as features to cluster the fracture models, the results shows that these model parameters may not be substantial to represent whole features for each fracture model. Hence model parameters clustering robust optimization has poor result compared with another methods. A more reasonable way to extract features from the model it to divide the model into several sub regions and extract features from each sub region. With more

features extracted, the model parameters clustering method should have better performance.

2. Besides to maximize the NPV, the other important objective for operators to design optimal well controls is to minimize the uncertainty in procedure of field development. Thus multi-objective optimization framework is proposed to maximum the expected NPV and minimize the risk of field development. However, the multi-objective optimization workflow would increase the number of required simulation runs significantly. Hence the multi-objective optimization with responses clustering should be investigated to improve the efficiency of the multi-objective optimization algorithm.
3. Selecting appropriate inflation factors is a heuristic but critical step in the ES-MDA. On the one hand, the appropriately selected inflation factors could improve the quality of data matching and final results. On the anther hand, bad inflation rates could cause ensemble collapse and generate too many extreme values of fracture fields. So developing an adaptive ES-MDA algorithm to select inflation rates automatically should also be an important work for history matching in NFRs.

BIBLIOGRAPHY

Brouwer, D. R., and J. D. Jansen. "Dynamic optimization of water flooding with smart wells using optimal control theory." *European Petroleum Conference*. Society of Petroleum Engineers, 2002.

Centilmen, Aydin, T. Ertekin, and A. S. Grader. "Applications of neural networks in multiwell field development." *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, 1999.

Chen, Chaohui. *Adjoint-gradient-based production optimization with the augmented Lagrangian method*. 2011.

Conn, Andrew R., Nicholas IM Gould, and Ph L. Toint. *Trust region methods*. Vol. 1. SIAM, 2000.

Li, Gaoming, and Albert C. Reynolds. "Uncertainty quantification of reservoir performance predictions using a stochastic optimization algorithm."*Computational Geosciences* 15.3 (2011): 451-462.

Nocedal, Jorge, and Stephen Wright. *Numerical optimization*. Springer Science & Business Media, 2006.

Sarma, Pallav, Khalid Aziz, and Louis J. Durlofsky. "Implementation of adjoint solution for optimal control of smart wells." *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2005.

Zhao, Hui, et al. "Maximization of a dynamic quadratic interpolation model for production optimization."*SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2011.

Aanonsen, S I., G.Navdal, D.S. Oliver, A C Reynolds "The ensemble Kalman filter in reservoir engineering--a review." Spe Journal 14.03 (2009): 393-412.

Emerick, Alexandre A., and A C. Reynolds. "Ensemble smoother with multiple data assimilation." *Computers & Geosciences* 55 (2013): 3-15.

Jansen, J D, D R Brouwer, and S G. Douma. "Closed loop reservoir management." *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2009.

Chen, Yan, Dean S. Oliver, and Dongxiao Zhang. "Efficient ensemble-based closed-loop production optimization." *SPE Journal* 14.04 (2009): 634-645.

Forouzanfar, Fahim. Well-placement optimization. The University of Tulsa 2012.

Chen, Chaohui, Gaoming Li, and Albert Reynolds. "Robust constrained optimization of short-and long-term net present value for closed-loop reservoir management." *SPE Journal* 17.03 (2012): 849-864.

Cui, H.Wen, and M G Kelkar. "Automatic history matching of naturally fractured reservoirs and a case Study." *SPE Western Regional Meeting*. Society of Petroleum Engineers, 2005.

Gang, Tao, and Mohan Gajanan Kelkar. "Efficient history matching in naturally fractured reservoirs." *SPE/DOE Symposium on Improved Oil Recovery*. Society of Petroleum Engineers, 2006.

Li, R, A. C. Reynolds, and D. S. Oliver. "History matching of three-phase flow production data." *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2001.

Warren, J. E., and P. Jj Root. "The behavior of naturally fractured reservoirs." *Society of Petroleum Engineers Journal* 3.03 (1963): 245-255.

Bai, M, Derek E, and Jean C R. "Multiporosity/multipermeability approach to the simulation of naturally fractured reservoirs." *Water Resources Research* 29.6 (1993): 1621-1634.

Moinfar, Ali, et al. "Development of a novel and computationally-efficient discrete-fracture model to study IOR processes in naturally fractured reservoirs." *SPE Improved Oil Recovery Symposium*. Society of Petroleum Engineers, 2012.

Moinfar, Ali, et al. "Comparison of discrete-fracture and dual-permeability models for multiphase flow in naturally fractured reservoirs." *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, 2011.

Li, Liyong, and Seong H. Lee. "Efficient field-scale simulation of black oil in a naturally fractured reservoir through discrete fracture networks and homogenized media." *SPE Reservoir Evaluation & Engineering* 11.04 (2008): 750-758.

Kazemi, H., et al. "Numerical simulation of water-oil flow in naturally fractured reservoirs." *Society of Petroleum Engineers Journal* 16.06 (1976): 317-326.

Pruess, Karsten. "A practical method for modeling fluid and heat flow in fractured porous media." *Society of Petroleum Engineers Journal* 25.01 (1985): 14-26.

Paico, Danny Hubert Rojas. *A New Procedure for History Matching Naturally Fractured Reservoirs*. Diss. Stanford University, 2008.

Hoteit, Hussein, and Abbas Firoozabadi. "An efficient numerical model for incompressible two-phase flow in fractured media." *Advances in Water Resources* 31.6 (2008): 891-905

Chen, B., and Reynolds, A.C. 2015. Ensemble-Based Optimization of the Water-Alternating-Gas-Injection Process. SPE J

Jiang, Jiamin, and Rami M. Younis. "A multimechanistic multicontinuum model for simulating shale gas reservoir with complex fractured system."*Fuel* 161 (2015): 333-344.

Jiang, Jiamin, and Rami Younis. "Hybrid Coupled Discrete-Fracture/Matrix and Multicontinuum Models for Unconventional-Reservoir Simulation." SPE Journal (2015).

Younis, Rami Mustafa, et al. Modern advances in software and solution algorithms for reservoir simulation. Stanford University, 2011.

APPENDIX A

SPATIAL AND TEMPORAL COVARIANCE MATRIX

It is a common assumption to assume that the log permeability and the porosity fields obey the Gaussian distribution in the geostatistic scope. To represent spatial and temporal correlation, several different covariance models have been introduced. The following covariance models are most commonly used for modeling random Earth processes: the spherical model for which:

$$C(h) = \delta^2 \begin{cases} 1 - \frac{3h}{2a} + \frac{h^3}{2a^2} & for\ 0 \le h \le a \\ 0 & for\ h > a \end{cases} \quad \text{A.1}$$

and the exponential family of covariance function,

$$C(h) = \delta^2 \exp\left(-3\left(\frac{|h|}{a}\right)\right). \quad \text{A.2}$$

For both formulas, h is the distance between two spatial locations and can be positive or negative. a is the correlation range. δ is the standard deviation of certain variables.

In cell-centered reservoir simulation models, the reservoir volume is discretized, and the cells are indexed two different ways: the i, j, k index refers to the row, column, and layer, while $m = i + N_i(j-1) + N_iN_j(k-1)$ indexes the grids with a single variable that takes values from 1 to N grid. N_i is the number of gridblocks in a row, and N_j is the number of gridblocks in a column of the model. The following equation illustrates the process of vectorizing a two-dimensional 10×10 array into a one-dimensional 100 element vector.

$$\begin{bmatrix} Z_{1,1} & \cdots & Z_{1,10} \\ \vdots & \ddots & \vdots \\ Z_{10,1} & \cdots & Z_{10,10} \end{bmatrix} \rightarrow \begin{bmatrix} Z_1 & \cdots & Z_{10} \\ \vdots & \ddots & \vdots \\ Z_{91} & \cdots & Z_{100} \end{bmatrix} \quad \text{A.3}$$

Denote the expectation of the random property at each gridblock in this 100 cell model as $M = [m_1, m_2, m_3, \dots, m_{100}]$. The covariance of the properties on the grid is the expectation of production of $(Z - M)$ and its transpose$(Z - M)^T$:

$$\text{Cov(z, z)} = E\left[(Z - M)^T (Z - M)\right] =$$

$$\begin{bmatrix} E[(Z_1 - m_1)(Z_1 - m_1)] & \cdots & E[(Z_1 - m_1)(Z_{100} - m_{100})] \\ \vdots & \ddots & \vdots \\ E[(Z_{100} - m_{100})(Z_1 - m_1)] & \cdots & E[(Z_{100} - m_{100})(Z_{100} - m_{100})] \end{bmatrix}. \quad \text{A.4}$$

Because the covariance of two random variables Z_i and Z_j is $cov(Z_i, Z_j) = E[(Z_i - m_i)^T (Z_j - M_j)]$, the A.4 is equivalent to:

$$\text{cov(Z, Z)} = \begin{bmatrix} cov(Z_1, Z_1) & \cdots & cov(Z_1, Z_{100}) \\ \vdots & \ddots & \vdots \\ cov(Z_{100}, Z_1) & \cdots & cov(Z_{100}, Z_{100}) \end{bmatrix}. \quad \text{A.5}$$

The spatial covariance matrix is always symmetric. With the assumption of stationarity the spatial covariance is only a function of lag distance. If each gridblock is 1 unit by 1 unit square and the value of the random variables are sampled at the gridblock centers, the covariance can be written in term of the distance between grid centers:

$$\text{cov(Z, Z)} = \begin{bmatrix} cov(0) & \cdots & cov(9\sqrt{2}) \\ \vdots & \ddots & \vdots \\ cov(9\sqrt{2}) & \cdots & cov(0) \end{bmatrix}. \quad \text{A.6}$$

The time-dependent covariance matrix for well controls shown as follows:

$$\text{cov}_{\text{wells}} = \begin{bmatrix} cov_{well1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & cov_{welln} \end{bmatrix} \quad \text{A.7}$$

$$cov_{welli} = \begin{bmatrix} cov^i(T_1, T_1) & \cdots & cov^i(T_1, T_x) \\ \vdots & \ddots & \vdots \\ cov^i(T_1, T_x) & \cdots & cov^i(T_x, T_x) \end{bmatrix} \quad \text{A.8}$$

$$cov^{i}(T_i, T_j) = \delta^2 \begin{cases} 1 - \frac{3(T_i - T_j)}{2T} + \frac{(T_i - T_j)^3}{2T^2} & for\ 0 \le T_i - T_j \le T \\ 0 & for\ T_i - T_j > T \end{cases} . \qquad \text{A.9}$$

APPENDIX B

CLUSTERING RESULT