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

THE GRADUATE SCHOOL

HISTORY MATCHING, PREDICTION AND PRODUCTION OPTIMIZATION WITH A PHYSICS-BASED DATA-DRIVEN MODEL

by
Zhenyu Guo

A dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy
in the Discipline of Petroleum Engineering

The Graduate School
The University of Tulsa

2018

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

THE GRADUATE SCHOOL

HISTORY MATCHING, PREDICTION AND PRODUCTION OPTIMIZATION WITH A PHYSICS-BASED DATA-DRIVEN MODEL

by
Zhenyu Guo

A DISSERTATION

APPROVED FOR THE DISCIPLINE OF

PETROLEUM ENGINEERING

By Dissertation Committee

Albert C. Reynolds, Chair
Rami Younis
Mustafa Onur
Richard Redner

ABSTRACT

Zhenyu Guo (Doctor of Philosophy in Petroleum Engineering)

History Matching, Prediction and Production Optimization with a Physics-Based Data-Driven Model

Directed by Albert C. Reynolds

186 pp., Chapter 6: Conclusions

(681 words)

Assisted history matching and life-cycle production optimization are the two key components of closed-loop reservoir management, which are traditionally performed based on a multitude of runs of full-scale reservoir simulation models which incur high computational costs for large-scale problems. To reduce the computational cost spent on full-scale simulation runs when performing history matching and production optimization, we focus on developing a new data-driven model, Interwell Numerical Simulation Model with Front Tracking (INSIM-FT). INSIM-FT can be built without any prior knowledge of geological information of the target reservoir. Although the INSIM-FT model is developed from production data and requires no prior knowledge of rock property fields, it incorporates far more fundamental physics than that of the popular Capacitance-Resistance Model (CRM). INSIM-FT also represents a substantial improvement on an interwell numerical simulation model (INSIM) developed by Zhao et al. (2016). Specifically, we introduce a theoretically correct procedure to compute water saturation in INSIM-FT that generally gives more robust and accurate solutions than are obtained with INSIM where saturations are computed with an ad hoc method. In addition, unlike INSIM, INSIM-FT incorporates the parameters defining power-law relative permeability curves as additional history-matching parameters

so that prior knowledge of relative permeabilities is no longer required as is the case with INSIM. Also, we introduce imaginary wells and their associated interwell connections (stream tubes) to enable more potential flow paths in the INSIM-FT model than are used in INSIM. These additional flow paths enable INSIM-FT to honor more correctly the physics than is done with the original INSIM model. With these modifications, one expects that INSIM-FT will be more robust than INSIM, and via computational examples, we show that this is the case.

After developing INSIM-FT for a two-dimensional reservoir model, we extend INSIM-FT to full three-dimensional multi-layered reservoirs, where it is necessary to consider gravitational effects. The extended model, which is referred to as INSIM-FT-3D, can be used for history matching, prediction and production optimization for a three-dimensional reservoir under waterflooding. Compared to the original INSIM-FT model, INSIM-FT-3D replaces the original Riemann solver in INSIM-FT by a new Riemann solver based on a convex-hull method that enables the solution of the Buckley-Leverett problem with gravity, where a fractional flow function may have more than one inflection point. Secondly, unlike the original INSIM-FT model, which assumes all wells are vertical, the INSIM-FT-3D model allows for the inclusion of wells with arbitrary trajectories with multiple perforations. Third, INSIM-FT-3D applies Mitchell's best-candidate algorithm to automatically generate the imaginary wells that are evenly distributed in the reservoir given a set of prefixed actual well nodes and fourth INSIM-FT-3D utilizes our own modification of Delaunay triangulation to build the 3D connection map necessary to use the general INSIM-FT-3D formulation.

The ensemble-smoother with multiple data assimilation (ES-MDA) is used for history matching with INSIM-FT or INSIM-FT-3D. The history matching parameters for INSIM-FT and INSIM-FT-3D are similar and include the connection-based parameters and the parameters that define power-law permeabilities. In addition to the common parameters included with the two methods, the parameters that define the well indices for the wells with multiple perforations are included in the INSIM-FT-3D model. For production optimization with INSIM-FT and INSIM-FT-3D, ensemble-based optimization (EnOpt) is used. Because

initially the developed data-driven models only allow rate controls, pressure controls cannot be used for production optimization. We provide a procedure to estimate the values of well indices via the history matching process if bottom-hole pressure (BHP) data are available. Then, BHP can be specified at each control step in the INSIM-FT(-3D) model to allow optimization of the life-cycle net present value (NPV) of production, where the producer's BHPs at control steps are included in the optimization variables.

Computational results show that, history matching and production optimization performed with INSIM-FT two- and three-dimensional models are far more computationally efficient than are those performed with full-scale reservoir simulation models but still give a characterization of a reservoir under waterflooding, future predictions and estimates of the optimal NPV of production, that are similar to those obtained using computationally expensive full-scale reservoir simulation models.

ACKNOWLEDGEMENTS

I would like to express my sincerest gratitude to my advisor, Dr. Albert C. Reynolds, for his inspiration, encouragement, and insightful inputs for my research. His earnest manner and responsive attitude towards scientific research has been a perfect model for me to learn.

I would also like to thank my dissertation committee members, Dr. Mustafa Onur, Dr. Rami Younis and Dr. Richard Redner, for taking time to review my dissertation and providing me with valuable feedback and comments.

Finally, I would like to show my greatest appreciation to my parents Minxin Guo and Xiaolan Zhang, for their unconditional support and love. I extend special gratitude to my wife, Di Zhu, for her love, understanding and encouragement throughout my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Closed-loop reservoir management is a general decision-making framework which combines the procedures of assisted history matching and production optimization. Its ultimate objective is to reduce the reservoir uncertainty, predict the reservoir future performance and maximize the economic return. The entire procedure of closed-loop management involves alternating data assimilation (history matching) steps with life-cycle production optimization based on the most recently updated reservoir model(s), to plan the optimal production strategy for the future. Since closed-loop reservoir management usually requires a multitude number of forward runs, keeping the computational cost of the forward model as low as possible is desirable.

As one essential part of closed-loop reservoir management, assisted history matching is an inverse problem which aims to find the model parameters (reservoir variables) that can honor historical dynamic data while maintaining geological plausibility. Due to uncertainties in the model parameters and noise in the observations of dynamic data, it has become popular to generate an ensemble of history-matched models for the purpose of assessing the uncertainty in reservoir properties and future reservoir performance predictions. Assisted history matching with uncertainty quantification is usually performed by using conventional full-scale simulators. Depending on the scale of the reservoir simulation model, a single forward simulation run may take the order of one hour to one or more days to complete. Assisted history matching with uncertainty quantification can take several hundred forward simulation runs, which is computationally expensive.

Life-cycle production optimization is also a vital part of closed-loop reservoir management. In life-cycle production optimization, given a reservoir model (or models) obtained

from history matching, one applies an optimization algorithm to estimate the well controls at all control steps (time intervals) which maximize some cost or objective function such as net-present-value (NPV) or total oil production. Traditionally, production optimization is performed by using full-scale grid-based simulators to predict the new value of the cost function at each time the optimization variables (well controls in this work) are updated during the iterative optimization process. Depending on the optimization method utilized, the number of forward simulation runs required to achieve the optimal well controls may range from the order of one hundred to a few thousand. Similar to assisted history matching, this process is computationally expensive for large-scale reservoir simulation models.

Due to the high computational cost for closed-loop reservoir management that involves the procedures of assisted history matching and life-cycle production optimization with full-scale reservoir simulation models, other methods to undertake the same tasks but with far less computational costs are badly needed. One possibility is to use a data-driven model, which can be treated as a computationally efficient surrogate model to replace the traditional full-scale reservoir simulation model.

In this research, we develop and utilize a physics-based data-driven model as a computationally efficient surrogate model to perform history matching and production optimization.

1.1 Literature Review on History Matching

The first focus of this research is to reduce the computational cost of assisted history matching by using a data-driven model. Assisted history matching is a complicated ill-conditioned inverse problem, to which an infinite number of solutions may exist that all match the observations, when the number of observed data is far less than the number of uncertain history-matched parameters to be tuned. Based on the Bayesian point of view, these solutions are represented by a set of reservoir models that follow a posterior distribution conditioned to the observed data. Many papers (Oliver et al., 1996; Reynolds et al., 1999; Nævdal et al., 2002; Gu and Oliver, 2007; Li and Reynolds, 2009; Chen and Oliver, 2012;

Emerick and Reynolds, 2012, 2013a,b; Le et al., 2016) have focused on obtaining at least an approximation of the correct sampling from the posterior distribution, which can properly reflect the uncertainty in the model space and the uncertainty in future predicted reservoir performance. To achieve this goal, a set of history matched models instead of a single model should be obtained to approximate the posterior sampling. This process requires a large number of full-scale forward simulation runs and incurs a heavy computational cost if the simulation model is large scale.

In general, the history matching problem can be considered as an optimization problem where the objective function usually consists of model mismatch and data mismatch terms. One or multiple history-matched models can be found by applying different types of optimization algorithms such as gradient-based optimization algorithms (Li et al., 2003; Reynolds et al., 2004; Gao and Reynolds, 2006; Kahrobaei et al., 2013), model-based derivative-free optimization algorithms (Zhao et al., 2013), direct pattern search derivative-free optimization algorithms (Gao et al., 2016), stochastic derivative-free optimization algorithms (Gao et al., 2004), and their hybrid counterparts. Starting from an initial guess of the reservoir model, the optimization procedure iteratively updates the model until a local or a global minimum of objective function is found. Multiple history-matched models can be found by starting from different initial guesses.

When an adjoint-based gradient is available, the gradient-based optimization algorithms (Li et al., 2003; Reynolds et al., 2004; Gao and Reynolds, 2006; Kahrobaei et al., 2013) perform better than other optimization algorithms that do not use the adjoint gradient. One of the popular gradient-based optimization methods used for history matching is the Gauss-Newton method, where the Hessian matrix of the objective function can be evaluated analytically using the first-order derivatives of data (or the sensitivity matrix). Alternatively, one may apply a quasi-Newton method such as limited-memory Broyden-Fletcher-Goldfarb-Shanno (LBFGS) (Liu and Nocedal, 1989). As benchmarked by Gao et al. (2016), Gauss-Newton methods perform better than quasi-Newton methods for history matching or least-square problems.

Unfortunately, the adjoint information is not generally available with commercial reservoir simulators, which restricts the application of adjoint-based history-matching methods. To take advantage of a gradient-based optimizer, we can approximate the gradient of the objective function that is to be minimized or maximized in different ways. One possibility is to use the finite difference based on a one-sided perturbation of each variable (Nash and Sofer, 1996). Unlike the adjoint-based methods which compute the gradient based on a single forward reservoir simulation plus a single "backward in time equivalent reservoir simulation run" to solve the adjoint system, evaluating the gradient by the finite-difference scheme requires n_p+1 simulation runs where n_p is the number of history matching parameters, thus, finite-difference approximation of the gradient becomes computationally infeasible when the number of reservoir parameters is large. An alternative method is to construct a quadratic interpolation model by multivariate perturbation (Powell, 2004), but this is also computationally infeasible for large-scale problems. With the numerically estimated derivative information, gradient-based optimization methods can also be applied to perform assisted history matching.

On the other hand, ensemble-based methods, which only implicitly estimate the sensitivity matrix, are widely used to solve history matching problems. A well-known ensemble-based history matching method is the ensemble Kalman Filter (EnKF) introduced by Evensen (1994), which has been extensively applied in many different areas. Since the first application of EnKF to reservoir history matching(Nævdal et al., 2002), the investigations of EnKF for the oil industry history-matching problems have been increasing significantly. However, EnKF has a major shortcoming which can introduce inconsistencies between the updated reservoir model parameters and updated state variables (primary variables solved for by the simulator) because it is used to update model parameters and sates simultaneously. To overcome this disadvantage of EnKF and improve its history matching quality, Emerick and Reynolds (2012, 2013a,b) developed an iterative version of ensemble smoother (ES) (van Leeuwen and Evensen, 1996). The Emerick-Reynolds algorithm is referred to as ensemble smoother with multiple data assimilation (ES-MDA). ES-MDA generally has bet-

ter accuracy than EnKF with comparable computational cost. Further studies of ES-MDA have been extended into history matching the production data of highly-channelized non-Gaussian reservoirs (Le et al., 2016). In this research, ES-MDA is used to perform history matching based on running a data-driven model as the forward model.

1.2 Literature Review on Production Optimization

The second focus of this research is to reduce the computational cost for production optimization related to repeatedly running full-scale reservoir simulations, since most works on life-cycle production optimization (Sarma et al., 2005; van Essen et al., 2006; Kraaijevanger et al., 2007; Jansen et al., 2009; Chen et al., 2009; van Essen et al., 2009a; Chen, 2011; van Essen et al., 2011; Chen et al., 2012; Fonseca et al., 2013; Oliveira and Reynolds, 2014; Chen and Reynolds, 2016; Liu and Reynolds, 2016b,a; Chen et al., 2017; Chen and Reynolds, 2018) are based on using a full-scale finite-volume or finite difference reservoir simulator as the forward model to predict the performance of the true reservoir for a given set of controls. The well controls can be specified as either well flow rates or flowing bottom-hole pressures during the production optimization period. In practice, considering the difficulty and cost of operations, the length of each control step used in life-cycle production optimization typically is equal to one to six months.

Evidence shows that gradient-based methods with the analytical gradient computed by the adjoint method are very computationally efficient for large-scale production optimization problems (Zakirov et al., 1996; Brouwer and Jansen, 2004; Jansen et al., 2005; de Montleau et al., 2006; van Essen et al., 2009b; Chen, 2011; Chen et al., 2012; Yan and Reynolds, 2014). However, the adjoint method is not typically embedded in commercial reservoir simulators used in the oil industry. The addition of the adjoint method to commercial simulator requires modification of the source code of the reservoir simulator and would be a challenging task.

The unavailability of the adjoint method motivated the development of ensemble-based methods for closed-loop reservoir management. Compared to adjoint-based methods,

ensemble-based methods treat the forward model as a black box and hence are more flexible to use, without the need to access the source code of the forward model. The basic idea of the ensemble-based method is to approximate the gradient of the objective function for production optimization with respect to the control variables from a ensemble of perturbed well controls. Unlike the finite-difference method (Yan and Reynolds, 2014), ensemble-based method does not require one perturbation and forward model run for each and every control variable, which significantly reduces the computational cost on gradient estimate compared to computation of the gradient using finite-difference approximation.

Ensemble-based optimization (EnOpt) (Chen and Oliver, 2009) is a well known ensemble-based method for production optimization. In Chen and Oliver (2009), EnOpt was tested for a large-scale SPE benchmark case, the Brugge synthetic field, which was the largest and most complex test case on production optimization at that time.

Do and Reynolds (2013) put forward a slight modification to the EnOpt algorithm developed by Chen and Oliver (2009). Do and Reynolds (2013) theoretically show that for production optimization based on a single realization (reservoir model), the stochastic gradient computed by three algorithms have the same expected value and provide a first-order approximation of the preconditioning covariance matrix times the true gradient. The three procedures include the stochastic Gaussian search direction, the simplex method and the modified EnOpt proposed in Do and Reynolds (2013). In order to introduce more temporal smoothness of control variables and prevent the abrupt change of well controls in time series, the gradient approximated by the ensemble is often pre-multiplied by the same preconditioning covariance matrix.

Robust life-cycle production optimization (van Essen et al., 2009b; Chen and Oliver, 2009; Fonseca et al., 2015, 2016) pertains to the case where the reservoir model is uncertain with the uncertainty represented by a suite (ensemble) of reservoir models. The usual objective for robust production optimization is to maximize the expectation of NPV, which is represented by the average NPV over all the reservoir models. Compared to production optimization based on a single model, robust optimization requires a greater number of sim-

ulation runs to finish the optimization procedure. However, by considering the geological uncertainties of one reservoir, robust production optimization allows one to characterize the uncertainty in the optimized NPV which is estimated as the expected NPV. EnOpt (Chen and Oliver, 2009) is the first standard algorithm designed for robust production optimization. As an improvement of the EnOpt gradient when the geological uncertainty is significant, the stochastic simplex approximate gradient (StoSAG) is proposed by Fonseca et al. (2016). In Fonseca et al. (2016), it is shown that StoSAG gives a theoretically more accurate approximation of the true gradient than is obtained with EnOpt, when the geological uncertainty is significant. Case studies show that StoSAG generally provide a significantly higher value of expected NPV for robust optimization than is generated with standard EnOpt procedure.

In this research, whether we perform robust optimization or assume the reservoir model is known is immaterial to the methodology presented in our research, so we use only a single history-matched data-driven model for life-cycle production optimization with the modified EnOpt algorithm introduced by Do and Reynolds (2013).

1.3 Literature Review on Surrogate Model

With the increasing need to reduce the computational cost related to running a full-scale reservoir simulation model to perform closed-loop reservoir management, different surrogate models have been developed as the replacement models to speed up the forward runs. Another reason why the surrogate model is useful is that sometimes there is insufficient data to build a reservoir model or building such a model cannot be economically justified. In such case, we have to use an alternative modeling procedure to generate a surrogate model which does not require as much information as needed for generating a reservoir simulation model.

A common surrogate model widely used is based on proper orthogonal decomposition (POD) (Markovinovic et al., 2002; van Doren et al., 2006; Cardoso and Durlofsky, 2010; Gildin et al., 2013; He and Durlofsky, 2014). Essentially, POD projects the system, or states of the system, into a lower dimensional space spanned by a relatively small basis (principal components). History matching and production optimization can then be performed by

adjusting the coefficients in a linear combination of the basis vectors. These reduced-order models are then considered as proxy models for full-scale reservoir simulators, and significantly enhance the computational efficiency by reducing the computational time spent on forward simulation runs. However, the development of these reduced-order proxy models requires comprehensive geological modeling as the full order reservoir simulation model has to be run at least one time to construct the reduced-order model.

On the other hand, data-driven models, which require no knowledge of detailed geological information, can serve as the proxy simulation model for predictions where a specific data-driven model is obtained by determining its parameters by history matching production data. The development of one data-driven model requires a relatively small investment of time compared to that required to develop and history match a full-scale simulation model. The efforts in building comprehensive reservoir models and the expensive computational cost to run full-scale simulators can be avoided if data-driven models are used for assisted history matching. In the early stage, some correlation-based data-driven models (Heffer et al., 1997; Refunjol, 1996; Jansen and Kelkar, 1997) were developed to distinguish the interwell geomechanics and flow trends based on the Spearman correlation coefficients, which infer geological features based on a simple statistical analysis without consideration of reservoir properties. The application of these models has been somewhat limited because the statistical results can be easily influenced by data noise. Also, due to the nature of correlation-based methods, the future reservoir performance is not quantitatively predictable. Model-based data-driven models overcome the aforementioned shortcomings of correlation-based models. Because input model parameters of the model-based data-driven models obtained by history matching better reflect the geological features of a reservoir better than those obtained by correlation-based data-driven models, one expects they can provide far more accurate predictions of future reservoir performance than can be obtained with correlation-based data-driven models.

Perhaps the most widely used data-driven model is referred to as the capacitance resistance model (CRM) (Yousef et al., 2006); the name arises from the analogy of flow in porous

media to the flow of an electric charge in a resistor-capacitor circuit, where the compressibility and transmissibility, respectively, are analogous to capacitance and resistance. Inspired by the work of Albertoni and Lake (2003), the CRM estimates the interwell connectivity by a series of allocation factors. These allocation factors approximately equal the percentage of the total water injection rate at each injection well that flows to each producer connected to that injection well. In CRM, each water injector is connected to every producing well. Gentil (2005) extended the model of Albertoni and Lake (2003) by providing a more physical explanation of the allocation factor and putting forth an empirical power-law model to estimate oil cut. By introducing a model parameter called the time constant in CRM, Yousef et al. (2006) overcame the limitation of the Albertoni and Lake (2003) model which ignored reservoir compressibility. CRM can efficiently estimate and match the total production rate data by knowing the injection rate information and, if available, the bottom-hole pressure (BHP) data at producers. Recent work on CRM has focused on extending its applicability to multi-phase flow and more complex field situations. Lake et al. (2007) applied the empirical power-law model (Gentil, 2005) in CRM to estimate the oil-phase rate and perform history matching and life-cycle production optimization. Since then, more empirical fractional flow models for miscible and immiscible flows (Sayarpour, 2008; Weber, 2009; Nguyen, 2012; Cao et al., 2015) were developed for use with CRM in order to estimate and history match the rates of different phases.

One defect of CRM is that it assumes a constant productivity index for each producer during the entire production history, which is not strictly physically correct in a multi-phase flow system. For example, in water-flooding, as water is constantly injected into a reservoir, the productivity indices of different producers must change with time since saturation profiles at producers change with time. Recent work on CRM (Cao, 2014; Cao et al., 2014, 2015) mitigate this problem by only matching a certain period of production data, during which the productivity indices approximately remain constant. However, the partial match of historical data evidently degrades the accuracy of predications generated with CRM. More importantly, the allocation factors, which are the most important model

parameters in CRM, are assumed to be constant during the entire production life. This assumption is tenuous because allocation factors change with flow pattern and injection rates (Thiele and Batycky, 2006). Also, the fractional flow models integrated with CRM originate from empirical formulations, which provide only an approximation of the correct physics.

Lerlertpakdee et al. (2014) proposed a flow-network model, which represents reservoir flow by a coupled-network model in which each pair of wells is connected with a one-dimensional (1D) finite-difference reservoir-simulation model. Two sets of model parameters, absolute permeability and pore volume, are defined at each gridblock in each 1D flow-network. The coupled system is solved for pressures and saturations semi-implicitly, effectively by a set of one-dimensional reservoir simulation runs. To do production optimization, a robust training scheme was developed to simultaneously history match several sets of simulated data obtained by running a full-scale simulator with different sets of well controls, which were expected to cover the range of control variables that are expected to be encountered during the iterations of a life-cycle production-optimization algorithm. Subsequently the trained flow-network model is applied to production optimization. Compared with CRM, the Lerlertpakdee et al. (2014) model uses the true relative permeability curves, which is physically correct but requires a priori knowledge of the relative permeability curves. The discretization of 1D interwell connections makes the flow-network model act as a simplified grid-based full-scale simulator, yet may result in a large number of model parameters for a large-scale reservoir with a large number of wells.

Another data-driven model, interwell simulation model (INSIM), was developed by Zhao et al. (2015) to history match reservoir performance in a water-flooding reservoir. Somewhat similar to the model of Lerlertpakdee et al. (2014), INSIM approximates a reservoir by a set of one-dimensional interwell connective units. INSIM solves for water saturation along each connective unit based on the Buckley-Leverett equation, which is somewhat similar to how streamline/streamtube simulator solves water saturation along each streamline/streamtube, but streamlines/streamtubes are defined a priori in INSIM. However, in-

stead of discretizing the 1D interwell connections by a series of finite-difference gridblocks as in (Lerlertpakdee et al., 2014), INSIM only defines a pair of parameters for each connection, which significantly reduces the number of model parameters. Unlike CRM which uses empirical equations to fit the oil cut, INSIM uses the true fractional flow curve as function of water saturation so that the oil cut (or oil-phase) can be estimated during a forward INSIM run. Nevertheless, the ad hoc computation of water saturation along each connection used in INSIM restrains its robustness. For difficult problems, e.g. the flow in a predefined interwell connection changes its direction or the water fronts from two injectors pass through an intermediate well node to arrive at another well node, INSIM can result in unreliable estimates of the water and oil phase rates.

In addition to a data-driven model, machine learning techniques, can be transported into petroleum industry to speed up the process of history matching and production optimization (Guo et al., 2017a,b). Although machine learning includes a set of different branches, non-linear regression may be the machine learning technique that is potentially most useful in history matching and production optimization, as this method enables us to build a response surface proxy model as a computationally efficient forward model that can be used to replace a full-scale reservoir simulation model. A response surface proxy is a parameterized mathematical formulation that approximates the input/output relation of one target function (He et al., 2015). For instance, the reservoir simulator that maps a series of model parameters to a series of reservoir flow responses can be approximated by a set of response surface proxies. Naively, polynomial regression is a candidate for building a response surface proxy of a reservoir simulation model, although it is not suitable for highly non-linear input/output relations. Other methods to build response surface proxy models that have been investigated to assist in solving history matching problems include Kriging (Landa and Güyagüler, 2003) and spline interpolation (Castellini et al., 2010). Both of these methods have the problem that universally exists for all interpolation methods, i.e. data overfitting if the training outputs are corrupted with noise. For reservoir simulation problems, the training data come from the simulated flow responses, which are mingled with

numerical noise from inexact solutions of both linear and nonlinear solvers. In the machine learning area, many of the regression methods are good enough to alleviate data over-fitting but also give reasonably small predictive bias. One such machine learning method that has been widely used is called support vector regression (SVR) (Smola and Vapnik, 1997).

In this research, our focus is to develop a physics-based data-driven model to efficiently perform history matching and production optimization.

1.4 Research Objectives

The accomplishment of this research are as follows:

1. Develop a new physics-based data-driven model, which is referred to as interwell numerical simulation model with front tracking (INSIM-FT) for history matching, future performance prediction and reservoir characterization of waterflooding performance.

2. Present a new modification of INSIM-FT that makes it possible to use bottom-hole pressures as well controls (design variables) when optimizing the life-cycle NPV of production; investigate whether a history-matched INSIM-FT model can find a solution of the well control problem for waterflooding optimization that produces an optimal NPV estimate that is a reasonable approximation of the optimal NPV generated by using a detailed reservoir simulation model for production optimization.

3. Extend the INSIM-FT model to allow for three-dimensional multi-layer reservoirs with gravitational effects and allow wells with arbitrary trajectories and multiple perforations. This model is referred to as INSIM-FT-3D

4. Explore the applicability of INSIM-FT-3D for production optimization including a complex field-scale synthetic application.

1.5 Dissertation Outline

The dissertation contains six chapters that proceed as follows:

In Chapter 2, we present the formulation of the INSIM-FT data-driven model for two-dimensional reservoirs. Then, the novel features of INSIM-FT compared to the original INSIM, are introduced. These features or modifications include the addition of imaginary wells, enforcement of a total rate constraint and a more robust way of computing interwell connectivity. Next, the objective function and constraints for history matching with INSIM-FT are introduced.

In Chapter 3, we first introduce the waterflooding optimization problem and its associated objective function; then we introduce the extension of INSIM-FT that allows bottom-hole pressures to be used as well controls; next the optimization and validation procedure are explained and applied. Following that, some examples are tested to demonstrate the superiority of INSIM-FT over INSIM and CRM.

In Chapter 4, first we introduce the pressure equation for INSIM-FT-3D for wells with a single perforation; secondly, the pressure equation is extended to consider wells with more than one perforation; thirdly, we give the saturation equation for INSIM-FT-3D and introduce a convex-hull method to solve the saturation equation; fourthly, Mitchell's best-candidate algorithm is introduced to generate imaginary wells; fifthly, we show how to use Delaunay triangulation to generate a connection map for history matching with INSIM-FT-3D; and finally three synthetic examples including a field-scale synthetic reservoir model and one field example are presented.

In Chapter 5, the methodology for production optimization with INSIM-FT-3D is introduced. Then, two examples including a field-scale Brugge reservoir are presented to show that INSIM-FT-3D is applicable for optimizing a waterflood.

We summarize the main conclusions of this research in Chapter 6.

Five appendices are included which give more mathematical details of the techniques developed and/or used. Appendix A gives a rigorous derivation of the equations that comprise the INSIM-FT model with the details of the front-tracking procedure used to compute the water saturation within INSIM-FT given in Appendix B. Appendix C

gives the derivation of CRM formulation that we use in our in house CRM software. Appendix D introduces the Graham's scan algorithm that can compute the convex hull of a set of points. We use this algorithm to solve the one-dimensional Buckley-Leverett problem with gravitational effects, where the fractional flow curve may have more than one inflection point. Appendix E describes the Mitchell's best-candidate algorithm that we use to generate the imaginary wells for three-dimensional reservoirs using INSIM-FT-3D.

CHAPTER 2

INSIM-FT MODEL FOR HISTORY-MATCHING, PREDICTION AND CHARACTERIZATION OF WATERFLOODING PERFORMANCE

In this chapter, we develop and employ a new data-driven model for assisted history matching production data from a reservoir under waterflood and apply the history-matched model to predict future reservoir performance. The new model is referred to as INSIM-FT, which is an acronym for Interwell Numerical Simulation Model with Front Tracking. INSIM-FT represents a substantial improvement on an interwell numerical simulation model (INSIM) developed by Zhao et al. (2016). INSIM-FT, as well as its INSIM predecessor, is based on the same underlying material balance as the IMPES reservoir simulation model with the following differences: (1) Unlike IMPES, no gridblocks are used in INSIM-FT, and only nodes and connections between nodes are used. (2) In INSIM-FT, individual gridblock parameters and properties such as petrophysical and fluid properties and gridblock dimensions are not used. Instead, INSIM-FT uses only two sets of parameters for each connection, the volume and transmissibility of each interwell region along a connection between a pair of node neighbors. (3) Although the same saturation equation used in IMPES is used in INSIM-FT, in INSIM-FT, a front tracking (FT) procedure is used to compute the water saturation on a finer scale than is represented by distances between pairs of nodes. This FT procedure enables the computation of flow when water from two different injection wells (node) arrive at the same node. (4)to run INSIM-FT as a forward model, one would have to know proper values of the interwell volumes and transmissibilities. As the values of these parameters cannot be directly computed from common reservoir simulation input, feasible values of these inter-nodal transmissibilities and volumes must be estimated by history-matching production data before INSIM-FT can be used for reservoir characterization, future reservoir

performance prediction, or as a forward model to estimate the well controls that maximize the net present value of production over the remaining reservoir life.

The predecessor of INSIM-FT, INSIM has been shown to work well for many examples. However, its original formulation contains three defects, namely, (i) unlike CRM and the new INSIM-FT formulation presented here, INSIM requires a priori knowledge of the relative permeability curves so that the phase rates at producing wells can be computed during an INSIM run; (ii) INSIM generally uses a single connection between an injector and neighboring producers which may not always provide a sufficient number of flow paths to obtain sufficient accuracy in history-matching and prediction; (although the authors of Zhao et al. (2016) were aware of the potential benefits of adding additional connections, they did not carefully pursue this modification) and (iii), most importantly, there is a fundamental theoretical issue with INSIM, namely that the method it employs to calculate water saturation is theoretically flawed for complex cases where the saturation distribution along a connection between wells cannot be computed from the standard Riemann solution; in such situations, the INSIM procedure for calculating water saturation is ad hoc. In particular, the INSIM method for calculating water saturation is theoretically incorrect whenever a producer is converted to a water injection well and would also be incorrect if, as in INSIM-FT, we use imaginary wells to provide more flow paths for injected water, i.e., adding more flow paths in INSIM would invoke the incorrect ad hoc procedure for computing saturation more often. The new data-driven model considered here uses a theoretically correct front-tracking procedure to calculate water saturation, hence the name INSIM-FT.

In this chapter, we introduce a theoretically correct procedure to compute water saturation in the INSIM-FT model. Also, new features are introduced to remedy the other two potential deficiencies of INSIM: (1) the parameters defining power law relative permeability curves are included as additional history-matching parameters so that prior knowledge of relative permeabilities is no longer required and (2) we add imaginary wells and their associated interwell connections (stream tubes) to enable more potential flow paths. Intuitively, one expects that INSIM-FT will be more robust than INSIM because of these modifications.

Table 2.1: Overall comparison among CRM, INSIM and INSIM-FT.

	CRM	INSIM	INSIM-FT
Requires relative permeabilities known	No	Yes	No
Provides accurate saturations	No	No	Yes
Add imaginary wells/additional flow paths	No	No	Yes
Adherence to correct physics of flow	No	Partially	Largely
Approximates reservoir connectivity	Yes	Yes	Yes

We note, however, even if there exist cases where INSIM-FT does not significantly improve the quality of the history-match or future reservoir performance predictions obtained from INSIM, INSIM-FT would still be preferable to INSIM because its procedure for computing water saturation is theoretically sound whereas the saturation calculation procedure used in INSIM becomes ad hoc under the circumstances noted above. An overall comparison of the assumptions and abilities of INSIM-FT, INSIM and CRM is shown in Table 2.1. This table also applies when the interior point optimizer used in the original INSIM code is replaced by ES-MDA to obtain a more robust implementation referred to as "INSIM (ES-MDA)."

To compare the performance of INSIM-FT with INSIM and CRM, we consider two synthetic examples in this chapter. To test the ability of the methods to estimate reservoir connectivity and monitor waterfloods, we compare results generated with the methods with those obtained with FrontSim (Schlumberger, 2013b). We also test a field example using INSIM-FT to illustrate the practical applicability of the method.

2.1 Methodology

Like INSIM (Zhao et al., 2015, 2016), INSIM-FT is an interwell numerical simulation model which characterizes the reservoir as a set of 1D connective flow volumes (streamtubes) as shown in Fig. 2.1. In Fig. 2.1, each red circle represents the static bulk volume controlled by the well at its center and this volume does not change with time. However, due to the rock compressibility, the associated pore volume, which is denoted by $V_{p,i}$, is a function of the average pressure in this volume and thus can change with time. Similarly, the dark gray regions in Fig. 1 represent the static constant bulk volumes between well nodes. The pore volume associated with the bulk volume between well i and well j is denoted by $V_{p,i,j}$ and in

the INSIM and INSIM-FT models, we assume that any change in this volume from its initial value is entirely due to rock compressibility effects. The transmissibility of connection (i, j) is denoted by $T_{i,j}$. The set of $V_{p,i,j}$'s and $T_{i,j}$'s that characterize these interwell pore volumes are model parameters in INSIM and INSIM-FT. INSIM involves no other parameters but INSIM-FT also includes in its parameter set some of the parameters that define power-law relative permeability curves.

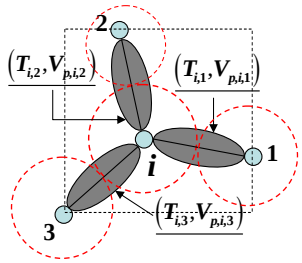


Figure 2.1: Connective units between wells of INSIM.

For completeness, in Appendix A, we provide a complete derivation of INSIM-FT, which partially but not entirely follows the work of Zhao et al. (2015). Oil field units are used throughout this Chapter. In INSIM and INSIM-FT, pressure equations are solved implicitly as in implicit-pressure-explicit-saturation (IMPES) to obtain the pressures at the well nodes. INSIM-FT replaces the INSIM procedure for calculating water saturation, which is not theoretically sound under all conditions, by a theoretically rigorous front-tracking procedure in order to calculate the saturation distribution along the interwell connection. The front tracking procedure used is presented in detail in Appendix B.

We define total mobility by

$$\lambda_t(S_w) = \frac{k_{ro}(S_w)}{\mu_o} + \frac{k_{rw}(S_w)}{\mu_w}, \tag{2.1}$$

where $\mu_m,\ m = o, w$ denote the oil and water viscosities in cp and $k_{rm},\ m = o, w$ represent the oil and water relative permeabilities. The transmissibility at well connection (i, j) is defined as

$$T_{i,j} = 1.127 \times 10^{-3} \frac{k_{i,j} A_{i,j} \lambda_t(S_{w,i,j})}{L_{i,j}}, \tag{2.2}$$

where the water saturation of connection (i,j), $S_{w,i,j}$, is decided by upstream weighting, i.e. if $p_i > p_j$, then $S_{w,i,j} = S_{w,i}$ with $S_{w,i}$ being the water saturation at well node i; otherwise $S_{w,i,j} = S_{w,j}$ with $S_{w,j}$ being the water saturation at well node j; $A_{i,j}$ is the average cross-sectional area in ft^2 of the flow connective unit between well i and well j; $L_{i,j}$ is the distance from well i to well j in ft; $k_{i,j}$ in md is the absolute permeability in the associated volume. The total compressibility at control volume i at time t^n is defined by

$$c_{t,i}^n = S_{o,i}^n c_o + S_{w,i}^n c_w + c_r, \tag{2.3}$$

where we assume throughout that the rock compressibility (c_r) and fluid compressibilities ($c_m,\ m = o, w$) in psi^{-1} are known constants; $S_{o,i}$ and $S_{w,i}$ are defined as the oil and water saturations on $V_{p,i}$ where $V_{p,i}$ is the pore volume corresponding to the volume associated with node i (see Fig. 2.1). From the IMPES formulation, the discrete form of the pressure equation for INSIM-FT, which is derived from a material balance on the control volume of well i, is given by

$$\sum_{j=1}^{n_{c,i}} T_{i,j}^{n-1}(p_j^n - p_i^n) + q_{t,i}^n = \frac{1}{5.615}\frac{c_{t,i}^{n-1} V_{p,i}^{n-1}}{\Delta t^n}(p_i^n - p_i^{n-1}), \tag{2.4}$$

where $n_{c,i}$ represents the number of wells that are connected to well i; pressures are in psi; the subscript n denotes the time step; $q_{t,i}^n$ in RB/day is the total rate of well i at t^n, where a positive value denotes injection and negative value denotes production, and the transmissibilities, ($T_{i,j}^{n-1}$'s) in Eq. 2.4 are defined as

$$T_{i,j}^{n-1} = 1.127 \times 10^{-3} k_{i,j} A_{i,j} \frac{\lambda_t(S_{w,i,j}^{n-1})}{L_{i,j}} = T_{i,j}^0 \frac{\lambda_t(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^0)}. \tag{2.5}$$

Throughout $\lambda_t(S_{w,i,j}^0)$ is specified as the total mobility at the initial water saturation, $S_{w,i,j}^0$, and $S_{w,i,j}^{n-1}$ is equal to the saturation at the upstream well node between the connection (i,j) at time level $n-1$, i.e. $S_{w,i,j}^{n-1} = S_{w,i}^{n-1}$ if $p_i^{n-1} > p_j^{n-1}$; otherwise $S_{w,i,j}^{n-1} = S_{w,j}^{n-1}$. In the examples,

we assume that the initial saturation is equal to the irreducible water saturation, S_{iw}, and that the relative permeabilities curves are obtained by normalizing effective permeabilities by the oil relative permeability at irreducible water saturations so that $k_{ro}(S_{iw}) = 1$. It follows that

$$\lambda_t(S^0_{w,i,j}) = \frac{1}{\mu_o}. \tag{2.6}$$

In the INSIM and INSIM-FT formulations, $V^{n-1}_{p,i}$ in Eq. 2.4 is assumed to satisfy the relation

$$V^{n-1}_{p,i} = 0.5 \sum_{j=1}^{n_{c,i}} V^{n-1}_{p,i,j}. \tag{2.7}$$

where $V^{n-1}_{p,i,j}$ is approximated by the first-order Taylor series expansion given by

$$V^{n-1}_{p,i,j} = V^0_{p,i,j}(1 + c_r(p^{n-1}_{i,j} - p^0)), \tag{2.8}$$

where p^0 is the initial reservoir pressure, and

$$p^{n-1}_{i,j} = 0.5(p^{n-1}_i + p^{n-1}_j). \tag{2.9}$$

With the assumption of Eq. 2.7, it follows that Eq. 2.4 involves only two types of parameters $T_{i,j}$ and $V_{p,i,j}$, which vary with time, with $T_{i,j}$ and $V_{p,i,j}$, respectively, given by Eqs. 2.5 and 2.8. INSIM assumes that relative permeabilities, viscosities, compressibilities and initial pressure are all known so that the only unknown parameters are the $T^0_{i,j}$'s and $V^0_{p,i,j}$'s. (INSIM-FT adds parameters defining power-law relative permeability functions as parameters.) According to Eq. 2.4, there is one pressure equation per well node. By solving the linear system of pressure equations, the pressure of each well control-volume at time t^n is obtained and the total liquid flow rate in RB/day along the connective volume (i, j) can be computed by

$$q^n_{t,i,j} = T^{n-1}_{i,j}(p^n_j - p^n_i), \tag{2.10}$$

where a positive value of $q_{t,i,j}$ means well j (or node j) is upstream of well i (or node i) and a negative value means well i (or node i) is upstream of well j (or node j). For the calculation of water saturation, we assume incompressible flow, so the partial differential equation governing the saturation distribution for linear flow through one of the connective volumes between a pair of nodes is in the form of Buckley-Leverett equation given by

$$\frac{\partial S_w(x,t)}{\partial t} + \frac{5.615 q_{t,i,j}(t)}{\phi_{i,j} A_{i,j}} \frac{\partial f_w(x,t)}{\partial x} = 0 \quad \text{for} \quad 0 \le x \le L_{i,j}, \quad t^{n-1} \le t \le t^n; \tag{2.11}$$

where $q_{t,i,j}(t)$ is assumed equal to $q^n_{t,i,j}$ from t^{n-1} to t^n; f_w is the water fractional flow; $\phi_{i,j}$ is the average porosity of the connective volume between wells i and j, and the x direction, which varies from connection volume to connection volume, is in the direction of the line segment that connects well i to j. In this chapter, we neglect both gravitational and capillarity effects so that f_w, as a function of water saturation, is given by

$$f_w(S_w) = \frac{1}{1 + \frac{k_{ro}(S_w)\cdot\mu_w}{k_{rw}(S_w)\cdot\mu_o}}. \tag{2.12}$$

Using

$$\phi_{i,j} A_{i,j} = \frac{V_{p,i,j}}{L_{i,j}}, \tag{2.13}$$

Eq. 2.11 can be rewritten as

$$\frac{\partial S_w(x,t)}{\partial t} + \frac{5.615 q_{t,i,j} L_{i,j}}{V^n_{p,i,j}} \frac{\partial f_w(x,t)}{\partial x} = 0 \quad \text{for} \quad 0 \le x \le L_{i,j}, \quad t^{n-1} \le t \le t^n. \tag{2.14}$$

Eq. 2.14 is solved semi-analytically from t^{n-1} to t^n using the front tracking procedure of Appendix B assuming that $q_{t,i,j}(t) = q^n_{t,i,j}$ for $t^{n-1} \le t \le t^n$. The initial condition for Eq. 2.14 is the saturation distribution for $0 \le x \le L_{i,j}$ at t^{n-1} along the well connection (i,j) which is given by

$$S_w(x, t^{n-1}) = S^{n-1}_w(x), \quad \text{for} \quad 0 \le x \le L_{i,j}. \tag{2.15}$$

The boundary condition for Eq. 2.14 is the water saturation, $S^u_{w,i,j}$, at time t^{n-1}, at the upstream node (either $x = 0$ or $x = L_{i,j}$) for flow between well i and j and this upstream saturation is assumed to be constant from t^{n-1} to t^n. If the upstream node is a water injector, then $S^u_{w,i,j} = 1 - S_{or}$, where S_{or} is the residual oil saturation, and $f_w(S^u_{w,i,j}) = 1$ all times.

The saturation equation in Eq. 2.14 applies to each connected well pair. The computation of the saturation profile at t^n along each connection by solving Eq. 2.14 is independent of the solutions along the other connections. However, when injected water from different injectors flows into a common node, the saturations flow out of the common node can only be resolved using the front tracking solution of Appendix B. After the saturation computation, the saturation profile of each connection at time level n has the form of

$$S_w(x, t^n) = S^n_w(x), \quad \text{for} \quad 0 \leq x \leq L_{i,j} \tag{2.16}$$

along each pair of connected nodes.

2.1.1 Relative Permeability Functions

In INSIM-FT, we use power-law relative permeability functions but estimate major parameters in these functions as part of the ES-MDA history-matching procedure (Emerick and Reynolds, 2012, 2013a,b), whereas in INSIM the relative permeabilities are assumed known a priori. Moreover, INSIM-FT allows one to use a different set of relative permeability curves along each connection although our experience with this application is limited and our brief experiments with those features suggest it is not necessary or useful for history matching. In all cases, the relative permeability curves are given by Corey-type relations. Specifically,

$$k_{rw}(S_{wn}) = a \cdot S^{n_w}_{wn} \tag{2.17}$$

and

$$k_{ro}(S_{wn}) = S^{n_o}_{wn}, \tag{2.18}$$

where

$$S_{wn} = \frac{S_w - S_{iw}}{1 - S_{iw} - S_{or}}. \tag{2.19}$$

In our applications, we assume that S_{iw} and S_{or} are known accurately for simplicity, and a, n_w and n_o are included as parameters to be estimated by history matching. The fluid viscosities, compressibilities and initial conditions are assumed known and fixed.

2.1.2 Front Tracking

The problem defined by Eq. 2.14 is, in general, a Cauchy problem which usually has a non-uniform initial condition, and reduces to the standard Buckley-Leverett problem (Riemann problem) only if the water saturation at the boundary condition is $S^u_{w,i,j} = 1 - S_{or}$ at all times and the initial condition is uniformly the irreducible water saturation along an interwell connection. The difference and connection between a Cauchy problem and a Riemann problem can be seen in Appendix B. Specifically, in the application of INSIM-FT, the standard Buckley-Leverett solution applies whenever we need to solve the saturation profile along a single connection between an injector and a downstream well (a producer or an imaginary well). However, the standard Buckley-Leverett solution is the solution of a Riemann problem and when well A is connected to multiple upstream wells and at least one downstream well, well B, the saturation between well A and well B, which still must satisfy Eq. 2.14, can no longer be represented as a Riemann problem because the saturation at well A is continually changing with time after water breakthrough at well A. INSIM uses an ad hoc procedure to compute water saturation when then standard Buckley-Leverett solution does not apply. In our work, we provide a general procedure for computing saturations via the implementation of a front tracking method conceptually similar to the one proposed by Holden et al. (1988) to solve the Cauchy problem. This method divides the Cauchy problem into a set of Riemann problems, which have analytical solutions. To avoid computing curved shock paths and simplify the problem, the rarefaction waves of the analytical solutions of Riemann problems are approximated by a set of small shocks. By connecting all the solutions of sub-Riemann problems, we obtain the global well-posed solution for Eq. 2.14. This method

has the advantage of being unconditionally stable and introduces little numerical dispersion. A complete derivation of the procedure is given in Appendix B.

2.1.3 Water Saturation and Oil Production Rate at Well Nodes

The front tracking algorithm gives the saturation profile along any pair of well connections. However, when flow into (or out of) the well node occurs along multiple paths that are directly connected to the well node, we need a way to compute the node saturation. Specifically, if one well, W_i (a producer or an imaginary well introduced later), is directly connected to multiple wells, the water fractional flow of W_i is computed as the ratio of the sum of the inflow water rates to the sum of the inflow total rate, i.e.,

$$f_{w,i}^n = \sum_{j=1}^{N_{c,i}} \frac{q_{t,i,j}^n f_{w,i,j}^n}{q_{t,i,j}^n}, \tag{2.20}$$

where $N_{c,i}$ is the number of wells that are connected to W_i and are also upstream of W_i; $f_{w,i,j}^n$ is the water fractional flow measured at the downstream node i along connection (i,j) at time level n. It is important to note that, the water fractional flows measured at node i for different connections (i,j), $j = 1, N_{c,i}$, are typically different, since the water saturation computed at the downstream end of each connection varies from connection to connection. The saturation at W_i is obtained by inverting the fractional flow function, i.e.,

$$S_{w,i}^n = f_w^{-1}(f_{w,i}^n). \tag{2.21}$$

Accordingly, if W_i (or simply i) is a producing well, the oil production rate at W_i is estimated as

$$q_{o,i}^n = \sum_{j=1}^{N_{c,i}} q_{t,i,j}^n \cdot (1 - f_{w,i,j}^n). \tag{2.22}$$

2.1.4 Demonstration Case

Here, we provide an example to demonstrate that the original INSIM algorithm does not always provide an accurate saturation distribution when the Buckley-Leverett solution

does not apply along all connections at all times. The example is a simple four-node, three-connection reservoir with four wells. The geometry of the reservoir is shown is Fig. 2.2. Note that there are four well nodes and three connected well pairs. Although the example In Fig. 2.2 seems quite simple, this example was chosen because it is representative of part of the more complete connection map that we use for cases with far more wells as will be clear when we discuss Example 1. The reservoir has homogenous absolute permeability and porosity fields. The three connections have the same average cross-sectional areas of 1,000 ft^2. Other properties are defined in Table 2.2. Note that to let the reservoir rock and fluid nearly incompressible, we set the rock and fluid compressibilites to very small values, i.e., in an order of magnitude of -10 psi^{-1}. The two injectors, W_1 and W_2, both inject water at the same constant rate of 100 RB/D. W_3 is a dummy well that does not inject or produce and W_4 is producing at a constant rate equal to 200 RB/D. We run the example with Eclipse (Schlumberger, 2013a), INSIM-FT and INSIM and observe the water cut at W_4. For this example, we assume that the correct relative permeability curves are known a priori.

This is an example where the water fronts from two injectors meet at a node, i.e., the water fronts from W_1 and W_2 meet at W_3. In this situation, the analytical solution of the Buckley-Leverett equation does not give the correct saturation distribution between W_3 and W_4 after water breakthrough at W_3, thus, we expect that INSIM cannot give the correct saturation profile between W_3 and W_4. As shown in Fig. 2.3, the water-cut obtained with INSIM-FT is in good agreement with that obtained with Eclipse, however, INSIM cannot predict the correct breakthrough time and water cut for W_4, i.e., INSIM-FT provides a more accurate saturation prediction than does INSIM.

2.2 INSIM-FT versus INSIM

In addition to a more reliable procedure for calculating saturation at nodes than is available in the original INSIM model, and adding terms defining power law relative permeability curves as parameters, INSIM-FT includes three other features that make it more robust and flexible than INSIM. These enhancements are discussed in the three subsections

Table 2.2: Property of the three-connection reservoir.

Property	Values
Absolute permeability, md	1,000
Porosity	0.2
Oil compressibility, psi^{-1}	3.4×10^{-10}
Water compressibility, psi^{-1}	3.4×10^{-10}
Rock compressibility, psi^{-1}	4.3×10^{-10}
Water viscosity, cp	1
Oil viscosity, cp	20
Initial reservoir pressure, psi	3,675

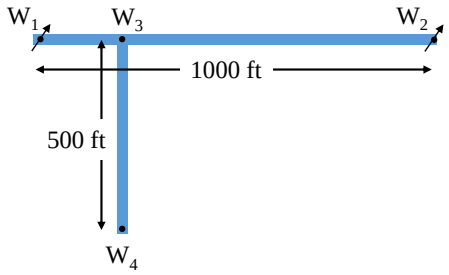


Figure 2.2: The geometry for a three-connection reservoir.

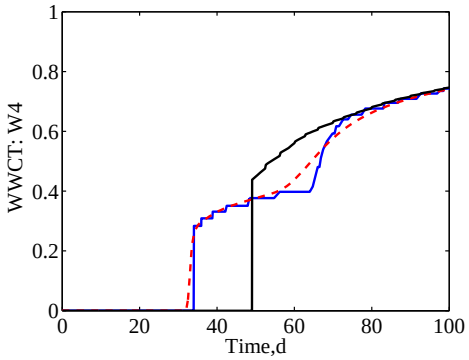


Figure 2.3: Water cut in W_4. The black line denotes the results obtained from INSIM, the blue line denotes the results obtained from INSIM-FT and the red dashed line denotes the results obtained from ECLIPSE.

below.

2.2.1 Imaginary Wells

In INSIM, all nodes represented the locations of actual wells and the only connections were between existing well pairs. In INSIM-FT, it is possible to add imaginary wells and associated connections in order to increase the possible flow paths along which injected

fluid can travel. INSIM-FT approximates a reservoir by a set of 1D connective units along which conservation equations are solved, which is similar to streamtube or streamline simulators. Thus, increasing the number of connections by adding imaginary well nodes effectively increases the number of streamtubes or streamlines along which fluid can be transported.

Another benefit of adding imaginary wells is that this avoids the direct connection between two producers or two injectors as done in Zhao et al. (2015). In INSIM, the flow pattern is always 1D and the flow direction between a connected well pair is from the well node with high pressure to the well node with low pressure, no matter the well type. Thus, if there is a direct connection of two injection wells, one of the injectors must serve as a sink, which is not consistent with the physics. Similarly, if two producers are connected, one of the producers most serve as a source whereas both producers should be sinks. Adding at least one imaginary well between each producer pair and each injector pair avoids the aforementioned issue. It is important to note the total flow rate at any imaginary well is zero, i.e., no fluid is produced or injected via an imaginary well.

It may be worthwhile to note that the aforementioned problem of one-directional flow between a pair of producers or a pair of injectors could seemingly be eliminated by deleting all direct connections between producer pairs and injector pairs but such a modification potentially eliminates the flow path or stream tube that enables the production of oil (or displacement of oil) from the interwell volume directly between the producer pair (injector pair). Thus we do not generally recommend the elimination of such connections.

The basic guidelines for adding imaginary wells are as follows: (1) At least one imaginary well must be added along a direct connection (stream tube) between two producers (or two injectors) as otherwise there is no way to accurately model the physics of flow between the two wells. For example, without adding an imaginary well between two producers, one producer must be upstream of the other and flow in the interwell region would flow from the upstream well to the downstream well whereas, in reality, part of the fluid between the two producers would flow to each producer. (2) There should be a sufficient number of imaginary wells added so that there are at least two distinct flow paths between each injector-producer

pair. These guidelines have proved satisfactory for the examples presented in this chapter including the field example as well as other synthetic examples we have tried. However, there is no guarantee that examples will not arise where the user may need to increase flow paths between injectors and producers to obtain an acceptable history match.

2.2.2 Total Rate Constraint on a Well's Control Volume

Another feature introduced in INSIM-FT is a rate constraint that states that the sum of the liquid rates flowing into the volume controlled by each producer is equal to the specified total rate of production from that producer. In the case where a well's control volume is very large, it may be possible for fluid to flow through the well's control volume without being produced by the well and in this case the rate constraint does not represent the true physics. Although our INSIM-FT code allows the user to choose whether or not to impose the rate constraint, all examples presented here impose this constraint because, via computational experiments not shown, we found that when the rate constraint is added to the optimization problem, we generally obtain better history matches than are obtained without the constraints.

If well i is a producer, the rate constraint becomes

$$q_{t,i}^n = \sum_{j=1}^{N_{c,i}} q_{t,i,j}^n, \tag{2.23}$$

where $q_{t,i,j}^n$ is computed from Eq. 2.10 and $q_{t,i}^n$ is the total liquid production rate specified at well i from t^{n-1} to t^n. Eq. 2.23 by itself is applied at each producing well. $N_{c,i}$ in the summation in Eq. 2.23 is over only connections that have a positive value of $q_{t,i,j}$ which is the case if and only if the producing well i is downstream of well j, i.e., flow is from node j into the control volume of well i. By enforcing Eq. 2.23 as a constraint when history matching, the outward flow rate from control volume i is approximately equal to $\sum_{j=1}^{N_{c,i}} q_{t,i,j}^n - q_{t,i}^n$, which is approximately zero. The details of how to enforce Eq. 2.23 will be discussed later.

2.2.3 Interwell Connectivity

For waterflood management, it is important to quantify which wells are "fed" by which injectors and one of the common ways to determine this is through the interwell connectivity. Interwell connectivity is defined in the capacitance resistance model (CRM) as an allocation factor $f_{i,j}$, which represents the percentage of the total water injection rate at each injection well i that flows towards producer j. Note that the subscripts i, j in CRM have different meanings than in INSIM-FT. These allocation factors are used to reflect large-scale geological characteristics. Unlike CRM, INSIM (Zhao et al., 2016) considers the allocation factor as a time-dependent variable so that the allocation factor for each injector-producer pair may vary from time step to time step in INSIM. Different from INSIM-FT, INSIM does not have imaginary wells in the connection map and there is only one streamtube between an injector and producer pair. For both INSIM and INSIM-FT, within the time interval $[t^{n-1}, t^n]$, the allocation factor $f_{i,j}(t)$ for the connection between injector j and producer i is assumed to constant and equal to $f_{i,j}^n$ with $f_{i,j}^n$ given by

$$f_{i,j}^n = \frac{q_{t,i,j}^n}{q_{t,j}^n}, \tag{2.24}$$

where $q_{t,i,j}^n$ is the total flow rate from j towards i and $q_{t,j}^n$ is the total injection rate of j at nth time step. It is important to note that $f_{i,j}^n$ represents an allocation factor, not the water fractional flow which is represented by $f_{w,i,j}$. For the purpose of comparison, in order to obtain a single-value presentation of the allocation factor for each injector-producer pair, the allocation factors in INSIM over all the INSIM steps are averaged by

$$\overline{f}_{i,j} = \frac{\sum_{n=1}^{n_t}(f_{i,j}^n q_{t,j}^n)}{q_{t,j}^n}. \tag{2.25}$$

In INSIM-FT, we provide a procedure to compute representative terms similar to the average allocation factors in INSIM that can reflect reservoir connectivity between injector-producer pairs. Specifically, we provide the means to compute the interwell total flow rate from an

injector that flows to a connected producer over the time interval $[t^{n-1}, t^n]$. The average total flow rate between an injector-producer pair can then be computed by averaging the total interwell flow rates calculated at the INSIM-FT time steps over the production history. When computing the connectivity, unlike INSIM, which only considers the flow due to the direct connection between an injector and producer, INSIM-FT also considers the flow from an injector that passes through intermediate nodes represented by imaginary wells to a producer.

To illustrate how to compute this "pass-through flow rate," we assume that the water from an injector will at most pass through one imaginary well before arriving at a producer. Consider the situation shown in Fig. 2.4, where injected water from I_2 must either directly pass through the interwell volume directly connecting the injector-producer pair, I_2 and P_1 or pass through either node Im_1 or Im_2, or more correctly, through the volumes associated with these imaginary wells before reaching P_1. We let q_{t,P_1,I_2} represents the total flow rate along the direct path connecting node I_2 to P_1, let q_{t,Im_1,I_j} for $j = 1, 2$ be the total flow rate between injector I_j and imaginary well Im_1 and let q_{t,P_1,Im_1,I_2} be the total flow rate from I_2 to P_1 that passes through the volume associated with imaginary node Im_1. Finally, we let q_{t,P_1,Im_1} be the total flow rate through the interwell volume directly connecting Im_1 and P_1. Then the flow rate from I_2 that passes through Im_1 and finally arrives P_1 is approximated by

$$q_{t,P_1,Im_1,I_2} = q_{P_1,Im_1} \cdot \frac{q_{t,Im_1,I_2}}{\sum_{i=1}^{2} q_{t,Im_1,I_i}}. \tag{2.26}$$

The pass-through flow rate from I_2 to P_1 through the volume associated with imaginary well Im_2 is denoted by q_{t,P_1,Im_2,I_2}. Then, the total interwell flow rate between I_2 and P_1, which is denoted by $\hat{q}_{t,P_1,I_2}$ and is given by

$$\hat{q}_{t,P_1,I_2} = q_{t,P_1,I_2} + \sum_{j=1}^{2} q_{t,P_1,Im_j,I_2}, \tag{2.27}$$

where $\hat{q}$ is used to indicate that this rate includes the pass-through rates plus the rate along

the direct connection between I_2 and P_j.

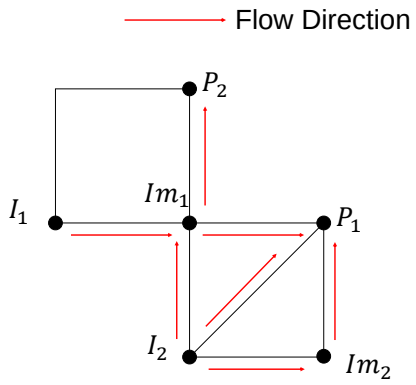


Figure 2.4: To compute the flow rate from I_2 passing through intermediate imaginary wells to P_1.

After the total interwell flow rate is obtained at every INSIM-FT time level, the average total interwell flow rate over the entire production history is used to represent the interwell connectivity. In the examples presented below, in order to validate the preceding procedure for determining how the water injection rate of a particular injection well is allocated among producers, we compare the connectivity obtained from INSIM-FT with that calculated from the true reservoir model with the FrontSim reservoir simulator of Schlumberger (Schlumberger, 2013b). FrontSim can calculate and output liquid flow rates between an injector-producer pair at every FrontSim simulation time level. By averaging these flow rates over the production history, we obtain the interwell connectivity in the same format as in INSIM-FT. Since FrontSim is run based on the same true geological model as Eclipse 100 which is used to generate the production data, the connectivity calculated with FrontSim reflects the connectivity of the true model. We also compare connectivity results obtained with FrontSim with those obtained with CRM and INSIM, where the allocation factors from CRM and INSIM are converted to the same format as used by FrontSim. Letting $\overline{q}_{t,i,j}$ be the connectivity used by INSIM-FT or FrontSim, we convert $\overline{f}_{i,j}$ in Eq. 2.25 to $\overline{q}_{t,i,j}$ by

$$\overline{q}_{t,i,j} = \overline{q}_{t,j}\overline{f}_{i,j}, \tag{2.28}$$

where $\overline{q}_{t,j}$ is the average total injection rate over the historical period. $f_{i,j}$ for CRM can be converted in a similar way.

2.3 History Matching

2.3.1 Parameters

In the original INSIM, we assume relative permeability curves are known, whereas in INSIM-FT, we assume relative permeabilities are of the form given by Eq. 2.17 and 2.18, where the exponents n_o and n_w and the coefficient a are unknown parameters to be determined by history matching. INSIM-FT includes the option of estimating one global set of relative permeabilities or estimating a different set along each interwell connection. The common parameters of INSIM and INSIM-FT are transmissibilities and pore volumes of connections at time zero. For $i = 1, 2, \ldots, N_{w-1}$, $j = i+1, i+2, \ldots, N_w$, $T^0_{i,j}$ represents the transmissibility of connection (i, j) at time zero and $V^0_{p,i,j}$ represents the pore volume of connection (i, j) at time zero. N_w is the number of wells which includes all the actual wells and imaginary wells. In practice, not every two wells are connected to each other and a connection map describing how wells are connected to each other will be given a priori for a reservoir before history matching. However, as discussed previously, we do not allow a direct connection between two injectors or between two producers in order to prevent a producer from becoming a source or an injector from becoming a sink. If well i and j are not connected in the connection map, then $T^0_{i,j}$ and $V^0_{p,i,j}$ will not be included as parameters for history matching. To satisfy the volume balance of the reservoir, the pore volumes of each connection at time zero must sum to the initial total pore volume of a reservoir, i.e.,

$$\sum_{i=1}^{N_w-1} \sum_{j=i+1}^{N_w} V^0_{p,i,j} = V^0_{p,\mathrm{tot}}, \tag{2.29}$$

where $V^0_{p,\mathrm{tot}}$ is the total pore volume of a reservoir at time zero and the two summations are over all the connected well pairs (i, j) in a predefined connection map. As an extension to INSIM, INSIM-FT has the option of whether to include $V^0_{p,\mathrm{tot}}$ as a parameter to be estimated

by history matching.

INSIM and INSIM-FT are both rate-control-based data-driven models, which means injection rates and total production rates must be specified at each time step. In INSIM and INSIM-FT, the observed oil production rates are the data that are history matched. Letting the vector d_{obs} be the vector of all observed oil production rates, the objective function that we minimize in INSIM and INSIM-FT to obtain a history match is given by

$$O(m) = \frac{1}{2}(g(m) - d_{\text{obs}})^T C_D^{-1}(g(m) - d_{\text{obs}}), \tag{2.30}$$

where C_D is the covariance matrix of measurement error, which is diagonal in all examples and d_{obs} is the vector of the observed oil production rates. In INSIM, m is the parameter vector that only includes $T^0_{i,j}$ and $V^0_{p,i,j}$ for all the connected well pairs, whereas in INSIM-FT a, n_o, n_w (see Eqs. 2.17-2.19) and possibly $V^0_{p,\text{tot}}$ are also included in the set of history-matched parameters. In our applications of INSIM-FT, the first two examples assume $V^0_{p,\text{tot}}$ is known but the field example uses $V^0_{p,\text{tot}}$ as a parameter. The constraints of the optimization problem for INSIM-FT are given by

$$\begin{cases} T^0_{i,j} \geq 0, \quad \text{for all connected well pairs } (i,j), & \text{(2.31a)} \\ 0 \leq V^0_{p,i,j} \leq V^0_{p,\text{tot}}, \quad \text{for all connected well pairs } (i,j), & \text{(2.31b)} \\ \displaystyle\sum_{j=1}^{N_{c,i}} q^n_{t,i,j}(m) = q^n_{t,i}, \quad \text{for } i = 1, 2, \ldots, N_{pro}, & \text{(2.31c)} \\ \displaystyle\sum_{i=1}^{N_w-1} \sum_{j=i+1}^{N_w} V^0_{p,i,j} = V^0_{p,\text{tot}}, \text{ where summations are over connected pairs } (i,j), & \text{(2.31d)} \\ 0 \leq a_{(i,j)} \leq 1, \quad \text{for all connected well pairs } (i,j), & \text{(2.31e)} \\ 1 \leq n_{o,(i,j)} \leq 6, \quad \text{for all connected well pairs } (i,j), & \text{(2.31f)} \\ 1 \leq n_{w,(i,j)} \leq 6, \quad \text{for all connected well pairs } (i,j), & \text{(2.31g)} \\ V^0_{p,\text{tot,low}} \leq V^0_{p,\text{tot}} \leq V^0_{p,\text{tot,up}}, \quad \text{if } V^0_{p,\text{tot}} \text{ is included as parameter}, & \text{(2.31h)} \end{cases}$$

where N_{pro} is the number of producers in the reservoir and $V^0_{p,\text{tot,low}}$ and $V^0_{p,\text{tot,up}}$, respectively, are the lower bound and upper bound for $V^0_{p,\text{tot}}$ which must be specified a priori.

In INSIM, the constraints for the history matching optimization problem (Eq. 2.30) only involve Eqs. 2.31a, b and d. Zhao et al. (2016) used an interior-point optimization algorithm in the Matlab optimization toolbox to solve this constrained optimization problem. However, the optimization tool from Matlab estimates the components of the gradient by finite-difference approximations, which can introduce errors as well as computational complexity if the number of parameters exceeds a few hundred. Here, we use the ensemble smoother with multiple data assimilation (ES-MDA) (Emerick and Reynolds, 2012, 2013a,b) for history matching with INSIM-FT. ES-MDA requires the generation of an ensemble of N_e initial models (realizations of m), m_k, $k = 1, 2, \ldots, N_e$, and these models actually provide a rank-deficient approximation of a prior covariance matrix which provides regularization during the history-matching process. ES-MDA effectively represents one iteration of the Gauss-Newton method with a matrix of average sensitivity coefficients for the case where the objective to be minimized is given by the $O(m)$ of Eq. 2.30 plus a regularization term (Reynolds et al., 2006).

In Eq. 2.31, Eqs. 2.31a,b,e-h only involve simple bound constraints; Eq. 2.31d is a linear equality constraint and Eq. 2.31c is a nonlinear equality constraint. First, we consider how to deal with the bound constraints and linear equality constraint. During the history matching procedure using ES-MDA, after each data assimilation step, the ensemble of reservoir models (m_k, $k = 1, 2, \ldots, N_e$) is updated using a update equation (Emerick and Reynolds, 2012, 2013a,b), which does not ensure that updated m_k satisfy the bound constraints or the linear equality constraint. Here, we propose to enforce the two types of constraints after each data assimilation step. For each updated reservoir model in the ensemble, the simple bound constraints are enforced by truncation, i.e., when an updated history-matching variable is greater than its upper bound, we set the variable equal to its upper bound, and whenever a variable is less than its lower bound, we set that the variable equal to its lower bound. The constraint of Eq. 2.31d should be enforced after Eq. 2.31c

is enforced, where Eq. 2.31c is enforced by history-matching ES-MDA update which will be discussed later. For one reservoir model m_k in the ensemble, we let $V^0_{p,i,j}$ obtained by enforcing Eq. 2.31c at each ES-MDA step be denoted by $V^{0,a,k}_{p,i,j}$, then we enforce Eq. 2.31d by replacing $V^{0,a,k}_{p,i,j}$ with $V^{0,a+,k}_{p,i,j}$, which is defined as

$$V^{0,a+,k}_{p,i,j} = V^{0,a,k}_{p,\text{tot}} \frac{V^{0,a,k}_{p,i,j}}{\sum_{i=1}^{N_w-1} \sum_{j=i+1}^{N_w} V^{0,a,k}_{p,i,j}}, \tag{2.32}$$

where $V^{0,a,k}_{p,\text{tot}}$ represents the ES-MDA updated $V^0_{p,\text{tot}}$ for kth ensemble member. So, Eq. 2.32 ensures that

$$\sum_{i=1}^{N_w-1} \sum_{j=i+1}^{N_w} V^{0,a+,k}_{p,i,j} = V^{0,a,k}_{p,\text{tot}},$$

for this specific realization of m_k. The procedure of Eq. 2.32 is repeated for all the other reservoir models at each ES-MDA step to ensure that Eq. 2.31d is satisfied for each ensemble member. Next, in order to deal with the nonlinear constraint of Eq. 2.31c, d_{obs} in Eq. 2.30 is expanded to include the specified total production rate as the observed data in addition to the observed oil production rate where the predicted total production rate at well i at t^n is given by Eq. 2.23, for $i = 1, 2, \ldots, N_{\text{pro}}$. C_D in Eq. 2.30 is also expanded to include the variance of the measurement error of the specified total rates, which must be specified a priori. $\sum_{j=1}^{N_{c,i}} q^n_{t,i,j}(m)$ is history matched to $q^n_{t,i}$. Notice that by doing the match of total production rate, $\sum_{j=1}^{N_{c,i}} q^n_{t,i,j}(m)$ is not guaranteed exactly equal to $q^n_{t,i}$ for each of the history matched reservoir models, because the history matching procedure cannot give the exact match; however, this match ensures that most inflow rate are produced from the producer. For all examples presented later, we obtain acceptable matches and predictions even though Eq. 2.31c is not exactly satisfied.

2.4 Application of INSIM-FT

To illustrate the accuracy of INSIM-FT, we consider three examples, the first example pertains to a homogeneous reservoir with a sealing fault, which was initially used by Cao et al. (2014) to show that CRM can identify the existence of fault. This example is included to

demonstrate that, similar to CRM, INSIM-FT can identify large-scale reservoir connectivity features such as a sealing fault. One can always raise the issue that our implementation of CRM is imperfect, but as in Cao et al. (2014, 2015), we use a in-house CRM implementation (Sayarpour, 2008). Our implementation of CRM is discussed in detail in Appendix C. In the second example, we consider a channelized reservoir and again show INSIM-FT outperforms CRM. The third example is a true field example where bottom water drive is the primary recovery mechanism. As there are no water injection well for this example, CRM is not directly applicable.

In order to obtain a quantitative measurement of the quality of our history matching results and future predictions for the methods considered, we define the normalized data mismatch for matching and predictions. The normalized data mismatch for a given model m is defined by

$$O_{N_d}(m) = \frac{1}{N_d}(g(m) - d_{\text{obs}})^T C_D^{-1}(g(m) - d_{\text{obs}}), \tag{2.33}$$

where for the data mismatch in the historical period, N_d is the number of observed oil production rates measured during the historical period, d_{obs} and C_D^{-1}, respectively, are the corresponding vector of observed data and the covariance matrix for measurement error; for the data mismatch in the prediction period, N_d, d_{obs} and C_D^{-1} in Eq. 2.33 are defined similarly for the prediction period. For history matching with INSIM and INSIM-FT, the average value of normalized data mismatches is computed over all the posterior realizations of the models.

2.4.1 Example 1: Homogeneous Reservoir with a Sealing Fault

We consider a reservoir with a sealing fault as shown in Fig. 2.5. Except for the fault, the reservoir has a homogeneous permeability field of 200 md and the porosity field is also uniform with ϕ=0.2. The reservoir has five injectors and four producers in a five-spot pattern. The completely sealing fault, shown in white as inactive gridblocks, divides this reservoir into two non-communicating regions. The performance of our in-house CRM and INSIM-FT software with/without adding imaginary wells are compared. To verify our implementation

of CRM, the same reservoir properties used by Cao et al. (2014) are input. The reservoir properties specified in Table 2.3 are the same properties used by Cao et al. (2014). The Eclipse reservoir simulation model representing the true model is based on a $33 \times 33 \times 1$ grid with grid dimensions being $\Delta x = \Delta y = 77.5$ ft and $\Delta z = 193.75$ ft. In the Eclipse reservoir simulation model, the five injectors are operated at a sequence of specified rates which change with time and four producers are produced at a constant bottom-hole pressures of 250 psi. The true data for history matching with INSIM-FT or CRM are obtained by running Eclipse 100 for 3,000 days with the true reservoir simulation model where the varying injection rates and constant producing BHP are specified as well operating conditions. CRM requires the injection rates to be specified as injection well operating conditions and only the total liquid well production rates are history matched. Then, the observed oil production rates for each well are fitted using an empirical equation. For history matching with CRM, the observed total liquid production rates and oil production rates, are obtained by adding uncorrelated Gaussian noise to the true data where the standard deviation of each measurement error is set equal to 2% of its true value. For history matching with INSIM-FT, the true total liquid production rates are specified as well operating conditions at each INSIM-FT time step and the observed data, the oil production rates, are the same ones as used in history matching with CRM. Note that, in INSIM-FT, since the sum of the total inflow rate into the control volume of one producer is required to be matched to the specified total production rate of that producer, we must also specify the standard deviation of the measurement error of each specified total production rate, which is specified as 2% of the specified total rate. Here, the data from the first 2,250 days of the Eclipse run are used for history matching and the remaining 750 days for future predictions.

The CRM-generated history match and future predictions of the total production are shown in Fig. 2.6. In this and similar figures presented later, the vertical line separates the historical and future-prediction time periods. The result of Fig. 2.6 is consistent with that from Cao et al. (2014), which suggests that our implementation of CRM is correct. It should be noted that in CRM, the oil-cut is fit by linear regression *after* history matching the

Table 2.3: Property of the fault-segmented reservoir

Property	Values
Oil compressibility, psi^{-1}	3×10^{-5}
Water compressibility, psi^{-1}	1×10^{-6}
Rock compressibility, psi^{-1}	1×10^{-6}
Water viscosity, cp	1
Oil viscosity, cp	1
Initial reservoir pressure, psi	1,250

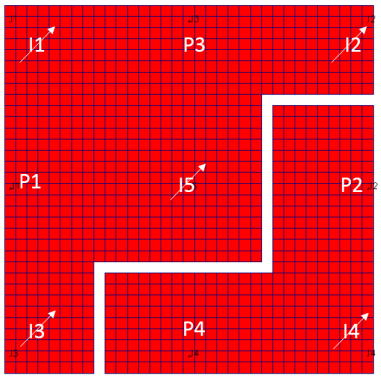
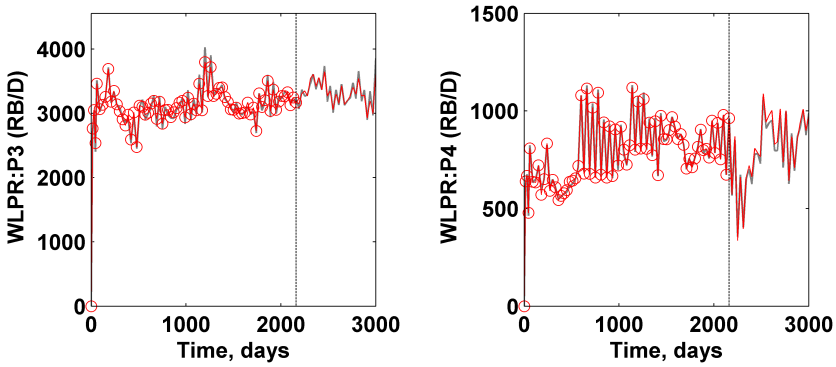


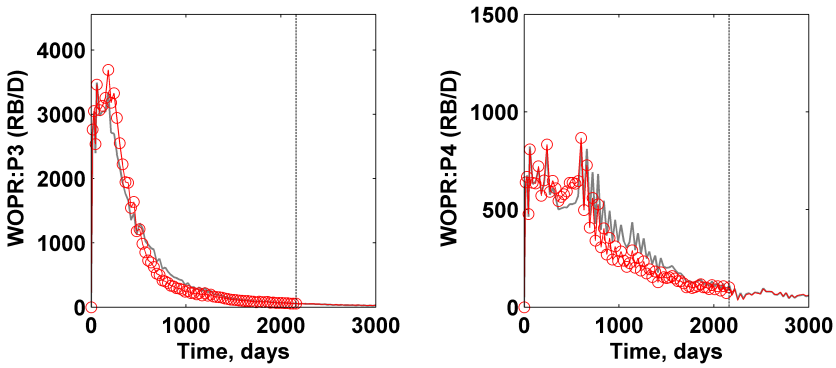
Figure 2.5: Fault geology.

total production rate and then the oil rate is calculated by multiplying the total production rate by the oil-cut, see Eq. C-25 in Appendix C; while in INSIM-FT, the oil production rate is directly computed and history-matched. Despite the empirical nature of the CRM calculation of the oil rate, Fig. 2.7 shows that the oil rate computed with CRM is in good agreement with the historical oil rate data and gives essentially a perfect prediction.



(a) CRM:P3 (b) CRM:P4

Figure 2.6: The total production rates obtained from the history matched CRM model. Red circles: observed total rates; red lines: true total production rates; gray lines: estimated total rates with the history-matched CRM model.



(a) CRM:P3 (b) CRM:P4

Figure 2.7: The oil production rates calculated from the history matched CRM model. Red circles: observed oil rates; red lines: true values of the oil rates; gray lines: estimated oil production rates by the history-matched CRM model.

INSIM-FT without Adding Imaginary Wells: The simplest possible well connections for INSIM-FT are shown in Fig. 2.8 for the reservoir of Fig. 2.5. The parameters required to be estimated include the $V^0_{p,i,j}$'s and $T^0_{i,j}$'s over all connections and a global set of parameters defining power-law relative permeabilities with $V^0_{p,\text{tot}}$ known a priori. The ensemble size for ES-MDA, N_e, is 250. It is assumed that the prior probability density function (PDF) for each parameter is given by a normal distribution. For each connected well pair (i,j), the mean value of $V^0_{p,i,j}$ for this normal distribution is estimated as

$$\overline{V}^0_{p,i,j} = \frac{L_{i,j}}{\sum_{k=1}^{N_w-1}\sum_{l=k+1}^{N_w} L_{k,l}} V^0_{p,\text{tot}}. \tag{2.34}$$

Then, each prior ensemble member, $V^{0,k}_{p,i,j}$ for $k = 1, 2, \ldots, N_e$ is randomly generated by sampling the Gaussian distribution $\mathcal{N}(\overline{V}^0_{p,i,j}, (0.2\overline{V}^0_{p,i,j})^2)$. According to Eq. 2.2, for each connected well pair (i,j),

$$T^0_{i,j} = 1.127 \times 10^{-3} \frac{k_{i,j} V^0_{p,i,j} \lambda_t(S^0_{w,i,j})}{\phi^0_{i,j} L^2_{i,j}}. \tag{2.35}$$

Replacing $\phi^0_{i,j}$ in Eq. 2.35 by ϕ^0, where ϕ^0 is the true porosity of the uniform porosity field, and substituting Eq. 2.6 into Eq. 2.35 yields

$$T^0_{i,j} = 1.127 \times 10^{-3} \frac{k_{i,j} V^0_{p,i,j}}{\mu_o \phi^0 L^2_{i,j}}. \tag{2.36}$$

Letting $\overline{T}^0_{i,j}$ be the mean value of $T^0_{i,j}$, $\overline{T}^0_{i,j}$ is estimated by replacing $k_{i,j}$ in Eq. 2.36 by $\bar{k}$ (the initial guess of the mean value of $k_{i,j}$) and is given by

$$\overline{T}^0_{i,j} = 1.127 \times 10^{-3} \frac{\bar{k} \overline{V}^0_{p,i,j}}{\mu_o \phi^0 L^2_{i,j}}, \tag{2.37}$$

where we use in Eq. 2.37 $\bar{k} = 500$ md, a value that is 150% higher than the true absolute permeability value of the uniform permeability field which is equal to 200 md; $\phi^0 = 0.2$ and $\mu_o = 1$ cp. For $k = 1, 2, \ldots, N_e$, the prior ensemble member, $T^{0,k}_{i,j}$, is generated by sampling $\mathcal{N}(\overline{T}^0_{i,j}, (0.2\overline{T}^0_{i,j})^2)$. The relative permeability model in this example is given by Eqs. 2.17-2.19 and the true values of the relative permeability parameters are given by

$$a_{\text{true}} = 0.6,\ n_{o,\text{true}} = 1.5,\ n_{w,\text{true}} = 1.5,\ S_{iw,\text{true}} = 0.2,\ S_{or,\text{true}} = 0.2.$$

To history match the relative permeabilities, S_{iw} and S_{or} are assumed to be known and the other three parameters are estimated with the history-matching procedure. The mean values of a, n_o and n_w are not equal to their true values, and, more specifically, are given as $\overline{a} = 0.653$, $\overline{n}_o = 1.70$, $\overline{n}_w = 1.46$. The prior PDF's of the three relative permeability parameters are specified by $\mathcal{N}(\overline{a}, 0.05^2)$, $\mathcal{N}(\overline{n}_o, 0.1^2)$, and $\mathcal{N}(\overline{n}_w, 0.1^2)$, where these PDF's must be also sampled to obtain the 250 initial realizations of relative permeability parameters. The oil production rates from multiple prior realizations of the INSIM-FT parameters are shown in Fig. 2.9. The history-match obtained with ES-MDA using INSIM-FT as the forward model is shown in Fig. 2.10. As can be seen, the match of oil production rate is not particularly good and in fact is much worse than the "match" of oil rates obtained with CRM. The prior and posterior ensemble of the oil-water relative permeability curves are shown in Figs. 2.11 and 2.12a respectively. The posterior oil-water relative permeability curves obtained by history matching are quite different from the true relative permeabilities. (However, as shown next

in 2.12b, much better history matches can be obtained if we introduce imaginary wells to provide additional connections and flow directions in the INSIM-FT model.)

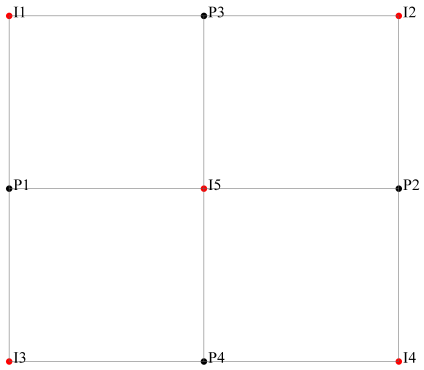
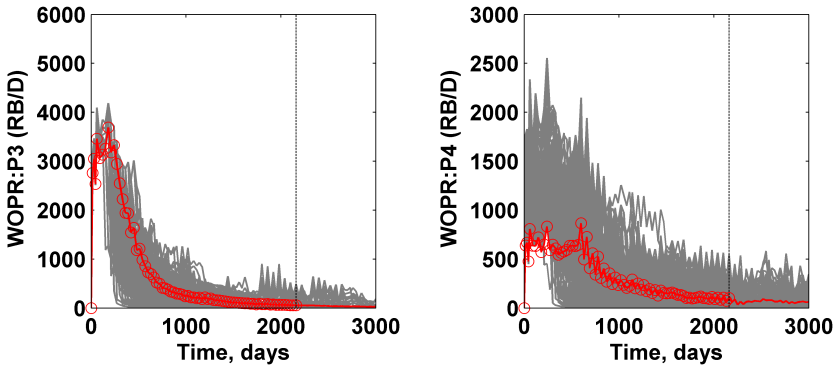


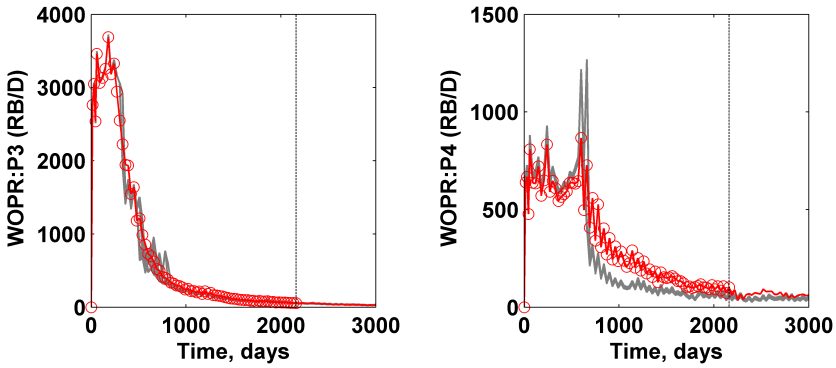
Figure 2.8: Connection map generated without adding imaginary wells.



(a) INSIM-FT:P3 (b) INSIM-FT:P4

Figure 2.9: The estimated oil production rates obtained from the prior INSIM-FT models. Red circles: observed oil production rates; red lines: true values of the oil production rates; gray lines: estimated oil rates obtained by running the prior INSIM-FT models; Example 1 without imaginary wells.

INSIM-FT with Imaginary Wells: Four imaginary wells are now added to this reservoir as shown in Fig. 2.13 with the new well connection map built based on interwell distances (Zhao et al., 2015) shown in Fig. 2.14. Note that the part of the full connection map of Fig. 2.14 that connects just the four wells I3, Im3, I5 and P4 is equivalent to the structure of the four well connection map of Fig. 2.2 provided that we equate wells I3, Im3, I5 and P4, respectively, in Fig. 2.14 to W1, W3, W2 and W4, respectively in Fig. 2.2.



(a) INSIM-FT:P3 (b) INSIM-FT:P4

Figure 2.10: The oil production rates obtained with the history-matched INSIM-FT models. Red circles: observed oil rates; red lines: true oil production rates; gray lines: estimated oil rates with the history-matched INSIM-FT models; Example 1 without imaginary wells.

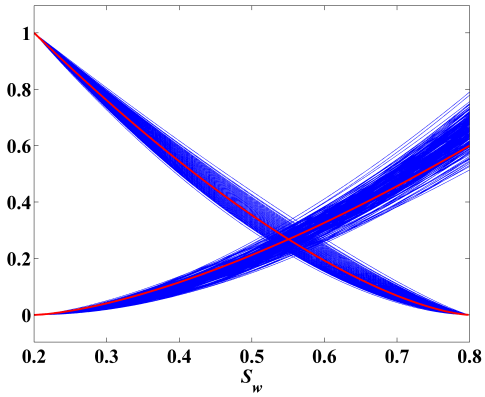
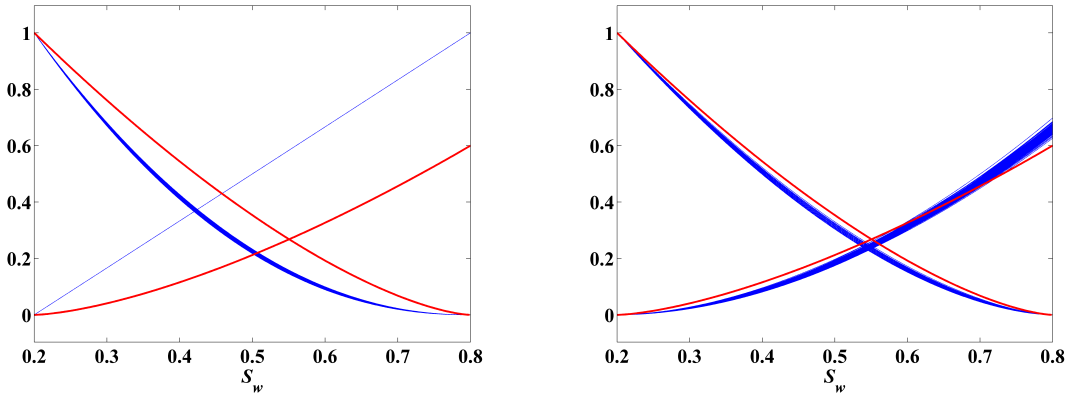


Figure 2.11: Prior oil-water relative permeability curves, fault. The red solid lines represent the true relative permeability curves and blue lines represent the prior models of relative permeabilities; Example 1 without imaginary wells.

The prior ensemble of vectors of model parameters is generated exactly the same way as for the case without imaginary wells. The oil production rates from multiple realizations of the vectors of model parameters based on the prior models are shown in Fig. 2.15. Fig. 2.16 shows the results from four methods where the INSIM (ES-MDA) results are generated by replacing the interior point optimizer used in the original INSIM code by ES-MDA. We made the change of history-matching algorithms because the finite-difference approximations used to compute the gradient for the interior-point optimizer sometimes yield a relatively poor gradient approximation, and because of this, the interior-point optimizer can fail to give a



(a) no imaginary wells

(b) imaginary wells added

Figure 2.12: Posterior oil-water relative permeability curves, fault. The red solid lines represent the true relative permeability curves and blue lines represent the history-matched models of relative permeabilities; Example 1 without imaginary wells.

good estimate of a minimizing set of model parameters. Throughout, the INSIM method obtained by replacing the interior-point optimizer by ES-MDA is referred to as INSIM (ES-MDA) whereas results labeled INSIM refer to results generated using the original INSIM code which implements a Matlab interior-point optimizer with all necessary derivatives computed by finite-difference approximations. It is important to note that imaginary wells are added only when INSIM-FT is used. The results of Fig. 2.16 illustrate that the history matches and predictions obtained from INSIM-FT, INSIM(ES-MDA) and CRM are fairly similar, whereas the INSIM results show poor matches and predictions. Table 2.4, however, provides a precise quantitative comparison of the history matching results from CRM, INSIM-FT, INSIM and INSIM (ES-MDA) based on the normalized data match. For this example and all examples that we have tried, INSIM (ES-MDA) gives a data match which is better than the one obtained with INSIM. For this example, the INSIM (ES-MDA) gives a substantially superior history-match than is obtained with INSIM but a worse history-match than is obtained with INSIM-FT. The results of Table 2.4 indicate that INSIM-FT produces the best history match and INSIM (ES-MDA) produces a better history match than CRM. The original INSIM algorithm produces a poor history match. It is important to note that the application of CRM gives a virtually exact prediction with the prediction error equal to 4.8

but this result is an abnormality because one should generally expect that the prediction error will be greater than or equal to the history-match error, a result that occurs with INSIM, INSIM (ES-MDA) and INSIM-FT as shown in Table 2.4. For further illustration, we reconsider the same example but use only the first 440 days of data when performing the history-match and use all of the other data to monitor prediction accuracy; the quantitative results are shown in Table 2.5. As the results indicate, the prediction error for CRM is no longer close to zero although it is still less than the prediction error obtained with INSIM-FT. However, CRM still yields a worse data match than is obtained with INSIM-FT or INSIM (ES-MDA).

Finally, by adding imaginary wells, we obtain a far more reasonable estimate of relative permeability curves, comparing the results of Fig. 2.12b where imaginary wells are added with the INSIM-FT estimates of Fig. 2.12a which uses the connection map of Fig. 2.8 with no imaginary wells added.

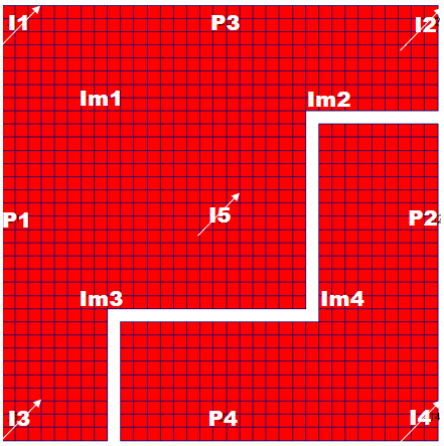


Figure 2.13: INSIM-FT well placement with adding imaginary wells; Example 1.

Table 2.4: Comparison of data mismatch for historical period (2250 days) and prediction period for CRM, INSIM, INSIM-FT and INSIM (ES-MDA) history-matched models; Example 1.

	CRM	INSIM	INSIM-FT	INSIM (ES-MDA)
Historical Period	74.3	752.2	54.5	62.2
Prediction Period	4.8	1202	89.4	85.2

A referee of the journal publication (Guo et al., 2018a) of some of the results of this

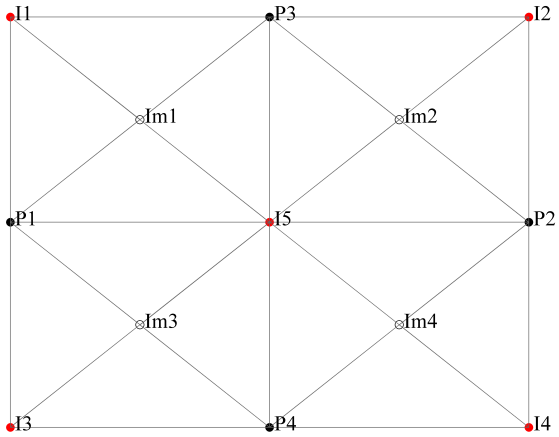
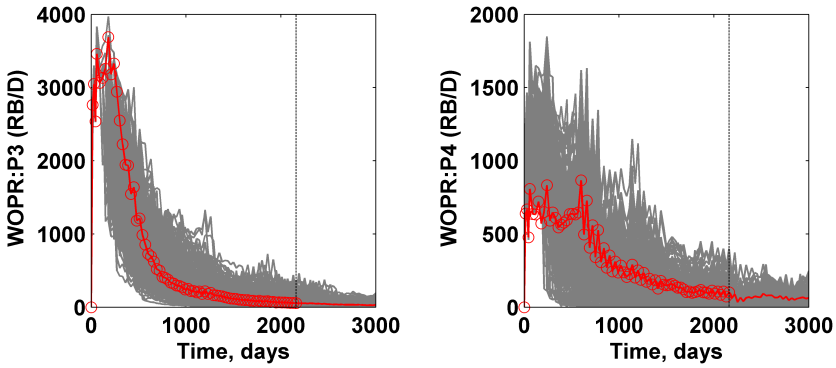


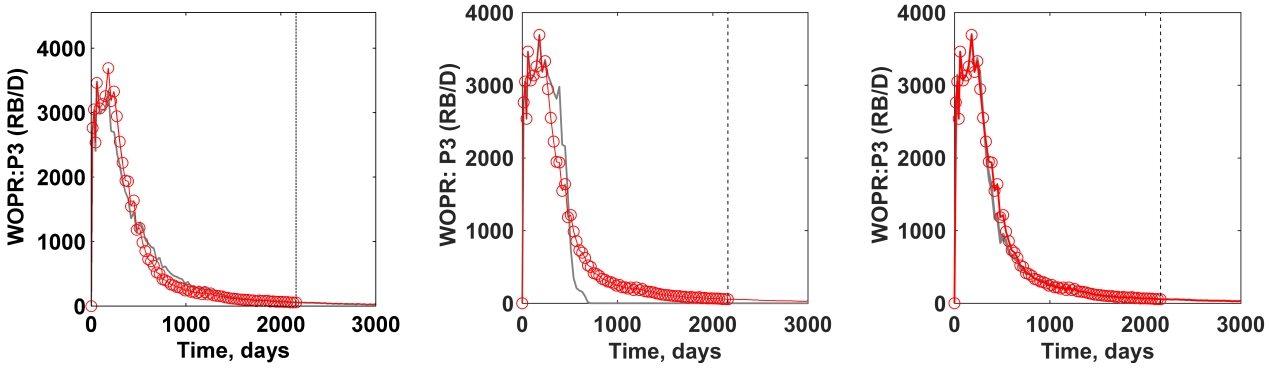
Figure 2.14: INSIM-FT connection map with imaginary wells and connections added; Example 1.



(a) INSIM-FT:P3 (b) INSIM-FT:P4

Figure 2.15: The oil production rates obtained from the prior INSIM-FT models. Red circles: observed oil rates; red lines: true oil rates; gray lines: estimated oil rates with the prior INSIM-FT models; Example 1 with imaginary wells.

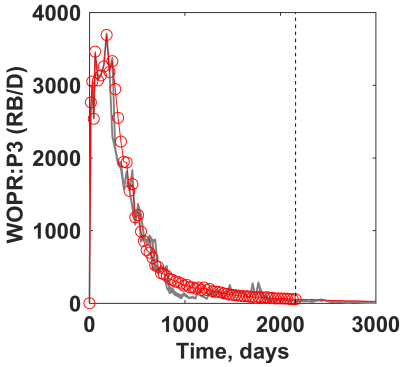
chapter posed a question about the overfitting of data so this topic may be worthy of a brief discussion. Because adding imaginary wells increases the number of parameters in INSIM-FT, the concern is that by continuing to add imaginary wells and hence connections and parameters, we will at some point encounter overfitting of data which corresponds to fitting noise, and when this happens, future predictions may become unreliable. To begin, we note that if m is Eq. 2.33 is set equal to the true model and d_{obs} is the random observation vector then the expected value of $O_{N_d}(m)$ is equal to unity; thus, the history-matching errors shown in Tables 2.4 and 2.5 indicate that we have not overfit the observed data. A theoretical basis for the expected values of O_{N_d} can be found in Tarantola (2005); also see the discussion



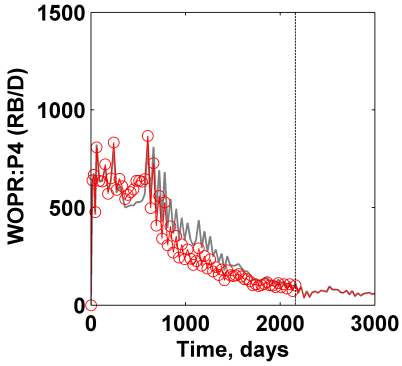
(a) CRM:P3

(b) INSIM:P3

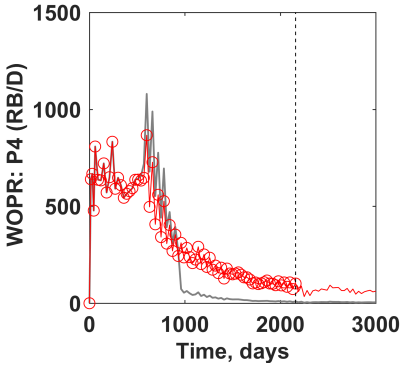
(c) INSIM (ES-MDA):P3



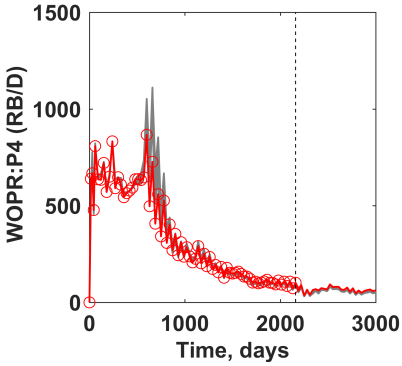
(d) INSIM-FT:P3



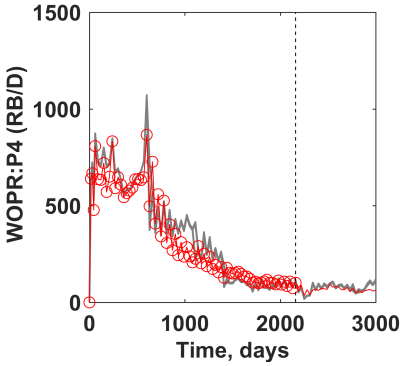
(e) CRM:P4



(f) INSIM:P4



(g) INSIM (ES-MDA):P4



(h) INSIM-FT:P4

Figure 2.16: Comparison of oil production rates calculated with the CRM, INSIM, INSIM (ES-MDA) and INSIM-FT history-matched models where only INSIM-FT uses imaginary wells; Example 1. Red circles: observed data; red lines: true data of oil production rate; gray curve: posterior oil production rate.

of the discrepancy principle in Vogel (2002) and Iglesias and Dawson (2013). Nevertheless, we cannot claim with certainty that if the number of interwell connections, and hence the number of parameters in INSIM-FT, were increased, the data would not be overfit, i.e.,

Table 2.5: Comparison of data mismatch for historical period (440 days) and prediction period for CRM, INSIM, INSIM-FT and INSIM (ES-MDA) history-matched models; Example 1.

	CRM	INSIM	INSIM-FT	INSIM (ES-MDA)
Historical Period	11.6	145.2	4.85	9.6
Prediction Period	237.3	3208.9	295	501

that the normalized objective function evaluated at the history-matched model would not be reduced below 1. In fact, the possibility of overfitting using ES-MDA with INSIM-FT may be more likely than with normal ES-MDA applications because there is no spatial correlation between the prior model parameters in the INSIM-FT and INSIM (ES-MDA) implementations. However, we should perhaps also note that adding imaginary wells is akin to adding streamlines and to the best of our knowledge, overfitting issues have not arisen when using streamline simulators in history matching even though there are parameters associated with each streamline.

Interwell Connectivity: In all INSIM algorithms, interwell connectivity is represented by the average interwell total flow rate between injector-producer pairs (see the discussion centered on Eqs. 2.26 and 2.27). Here, the accuracy of connectivities estimated with the INSIM, INSIM(ES-MDA), INSIM-FT and CRM algorithms is investigated by comparison with connectivity generated by running the FrontSim simulator using the true reservoir model as input. In FrontSim, the interwell total flow rate between a well pair can be output at each time step of the entire history-matched period. The average total flow rate is calculated as the average of all these output flow rates over the history-matched period. With CRM, however, the interwell connectivity is represented by allocation factors which are defined in Yousef et al. (2006). As INSIM (Zhao et al., 2015) and INSIM(ES-MDA) utilize a connectivity formulation which is equivalent to the allocation factor in CRM. Thus, to obtain a valid comparison, the allocation factor between each injector-producer well pair obtained with CRM, INSIM or INSIM (ES-MDA) is converted to the average interwell total flow rate by simply multiplying the allocation factor by the average total injection rate of the associated

injector over the production history.

For comparison, the interwell connectivities obtained from FronSim, CRM, INSIM, INSIM (ES-MDA) and INSIM-FT, respectively, are shown in Fig. 2.17a, 2.17b, 2.17c, 2.17d and 2.17e. In these and similar figures, the length of each narrow red triangle represents the magnitude of the interwell connectivity and the direction points to the producer associated with the injector-producer well pair. The number alongside each red triangle indicates the value of the average total interwell flow rate (over the history-matching period) between the injector-producer pair in units of RB/day. For brevity, the values of interwell flow rate are shown in the figure only for the connections with a large interwell flow rate. As shown in 2.17b, the existence of the fault is reflected by the fact that the interwell flow rate translated from the CRM allocation factor is small (on the order of 10 RB/day or less) between every injector-producer pair that is separated by the fault, i.e, lie on different sides of the fault. However, unlike the CRM result of Fig. 2.17b, the FrontSim result of Fig. 2.17a indicates that there is no flow between (I2, P1) and (I3, P3). In fact, the FrontSim streamline field, which is not shown here, indicates that there is no streamline from I3 to P3, or from I2 to P1. Thus, the CRM result showing a large flow rate from I3 to P3 and I2 to P1 seems to be incorrect. The INSIM results are unreasonable in the sense that there is a considerable amount of flow between injector-producer pairs that cross the sealing fault; for example, the average flow rate between injector I5 and P4 from the INSIM results is 265.6 STB/D as opposed to the correct value of zero which follows from the fact that the I5 and P4 are on opposite sides of the fault. Note the INSIM (ES-MDA) results are consistent with the presence of the fault, and, for this example are of the same accuracy as those obtained with INSIM-FT. The connectivity estimates obtained with CRM are clearly inferior to those obtained from INSIM-FT and INSIM (ES-MDA) based on the comparison with the FrontSim results.

2.4.2 Example 2: Channelized Reservoir

In this example, history matching is performed with INSIM-FT, INSIM, INSIM(ES-MDA) and CRM for a channelized reservoir. The log-permeability field is generated based

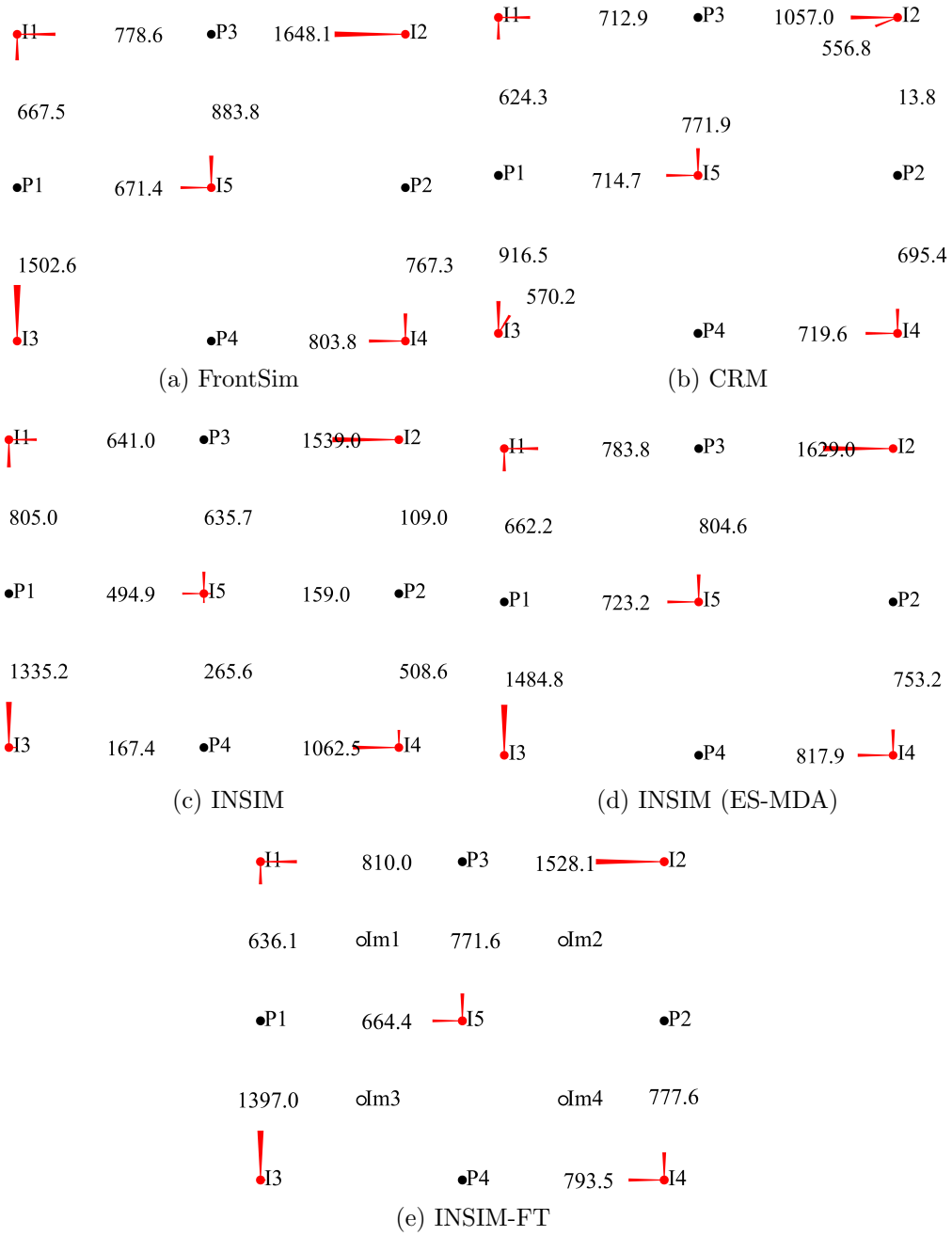


Figure 2.17: Interwell connectivity obtained from FrontSim, CRM, INSIM, INSIM (ES-MDA) and INSIM-FT for the fault case. The length of a narrow red triangular denotes the magnitude of total interwell flow rate between an injector-producer pair and the direction of the triangular indicates which producer belongs to this injector-producer well pair.

on a $25 \times 25 \times 1$ geological grid which is shown in Fig. 2.18. The red/yellow regions represent high permeability channels. However, the reservoir simulation model is based on the refined grid obtained by subdividing each geological grid block into a $3 \times 3 \times 1$ sub-grid with each of

the 9 subgrid blocks having the same value of permeability as the host coarse grid cell. Thus the reservoir simulation model, which represents the true reservoir is based on a $225 \times 225 \times 1$ grid where the size of each of these gridblock is 90 ft by 90 ft by 10 ft where 10 ft is the thickness of the gridblock. The reservoir has a homogeneous porosity field with $\phi = 0.2$. Other reservoir properties are shown in Table 2.6. The well locations shown in Fig. 2.18 indicate that we have introduced twelve imaginary wells in order to increase the pathways along which water can flow from injectors to producers. The resulting connection map for INSIM-FT with imaginary wells is shown in Fig 2.19. We use 800 days of production data for history matching and then predict reservoir performance for an additional 200 days. The true total production rates and oil production rates are obtained by running the $225 \times 225 \times 1$ Eclipse model with the well operating schedule defined in Table 2.7. For history matching with CRM, the total production rates are the observed data to be history matched and the oil production rates are predicted with the history matched CRM model. The corresponding observed total and oil production rates for matching with CRM are obtained by adding uncorrelated Gaussian noise with standard deviation equal to 2% of the true data. However, for history matching with INSIM-FT, the true total production rates are specified as well controls and the observed oil production rates (the same data used for matching with CRM) are history matched. In order to match the sum of the inflow rate into the control volume of one producer to the total production rate of that producer with ES-MDA, the standard deviation of the measurement error of each specified total production rate is set to 2% of the specified rate.

Table 2.6: Property of the channelized reservoir

Property Name	Values
Porosity	0.2
Oil compressibility, psi^{-1}	3.4×10^{-5}
Water compressibility, psi^{-1}	3.4×10^{-5}
Rock compressibility, psi^{-1}	4.3×10^{-5}
Water viscosity, cp	1
Oil viscosity, cp	20
Initial reservoir pressure, psi	3,675

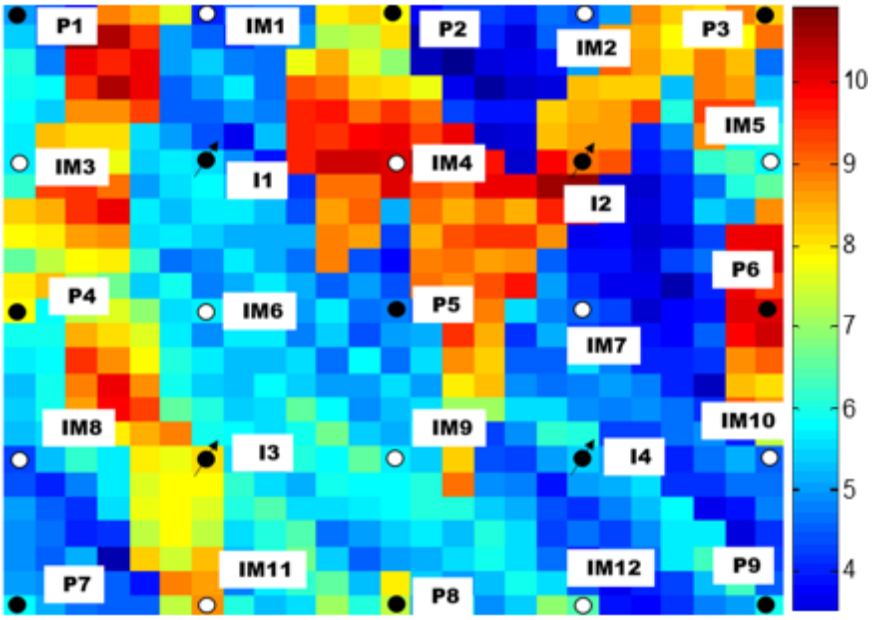


Figure 2.18: Log permeability field, channelized reservoir.

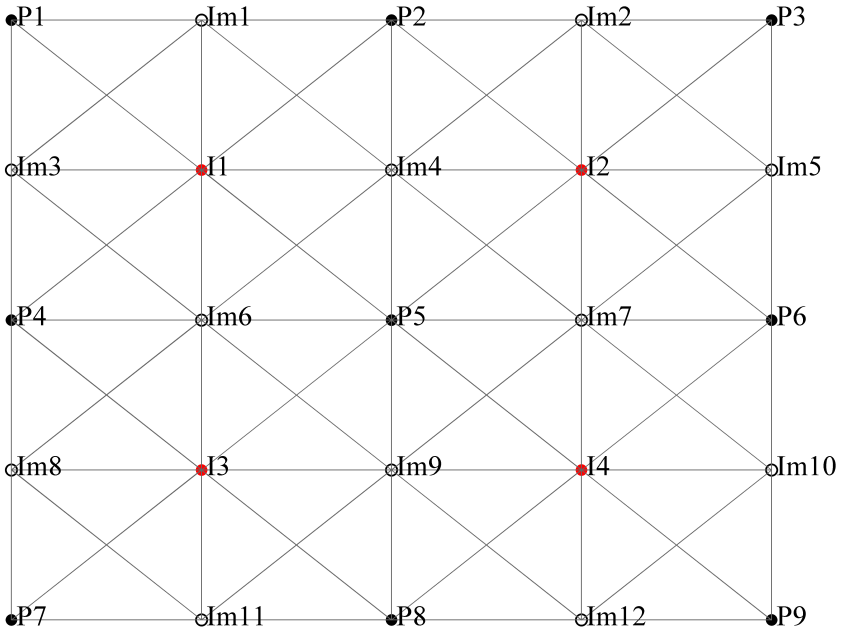
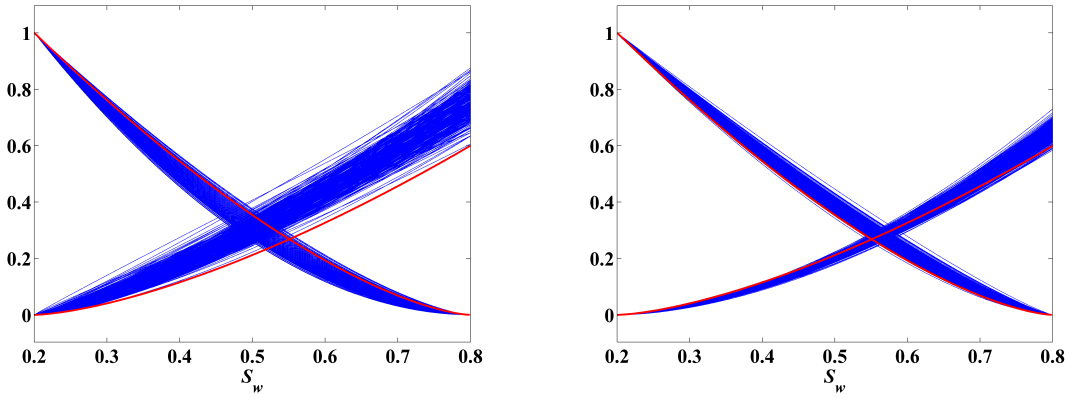


Figure 2.19: Connection Map, channelized reservoir.

For ES-MDA, the initial ensemble of model parameters includes initial guesses for transmissibilities and pore volumes at time zero defined on connections and parameters defining power-law relative permeabilities. $V^0_{p,\text{tot}}$ is known a priori in this example. The ensemble size, N_e is equal to 300. For all the connected well pair (i, j), the kth prior realization of $V^{k,0}_{p,i,j}$ is generated by sampling the normal distribution $\mathcal{N}(\overline{V}^0_{p,i,j}, (0.2\overline{V}_{p,i,j})^2)$

with $\overline{V}^0_{p,i,j}$ given by Eq. 2.34. Similarly, for $k = 1, 2, \ldots, N_e$, the prior $T^{k,0}_{i,j}$ is sampled from the normal distribution $\mathcal{N}(\overline{T}^0_{i,j}, (0.2\overline{T}^0_{i,j})^2)$ with $\overline{T}^0_{i,j}$ defined in Eq. 2.37, where $\bar{k} = 1000$ md is a random guess of the average absolute permeability of the entire reservoir; $\mu_o = 20$ cp and $\phi^0 = 0.2$ are true values. Here, we use the same true relative permeability curves as in the first example, but the three relative permeability parameters are distributed as $a \sim \mathcal{N}(\overline{a}, 0.05^2)$, $n_w \sim \mathcal{N}(\overline{n}_w, 0.1^2)$ and $n_o \sim \mathcal{N}(\overline{n}_o, 0.1^2)$ with the mean values given by $\overline{a} = 0.75$, $\overline{n}_w = 1.3$, $\overline{n}_o = 1.7$. Again, the values of S_{iw} and S_{or} are assumed to be known. We assume a single global set of relative permeabilities which are shared by all volume connections.

In this example, an initial ensemble of 300 realizations of the INSIM-FT model parameters is generated by sampling the aforementioned Gaussian distributions. The prior and history-matched realizations of relative permeabilities are shown in Figs. 2.20a and 2.20b, respectively. The oil production rates calculated with INSIM-FT using the multiple prior realizations of model parameters are shown in Fig. 2.21. Fig. 2.22 compares oil production rates calculated with the CRM, INSIM, INSIM(ES-MDA) and INSIM-FT history-matched models for four wells, namely P1 to P4. In this case, the INSIM and INSIM(ES-MDA) algorithms give a good history match of oil rate data at wells P1, but the relatively poor data match and predictions obtained for wells P3 and P4 are more representative of the overall INSIM and INSIM(ES-MDA) results as indicated by the computational results shown in Table 2.8. In Table 2.8, we see that for this example, the INSIM (ES-MDA) data matches are significantly worse than those obtained with INSIM-FT. In fact, unlike the Example 1 results, in this example, the INSIM (ES-MDA) history match is even inferior to the one obtained with CRM. Table 2.9 presents results with the same example but with the history-matching period shortened to 400 days. It is important to note that the results of Tables 2.8 and 2.9 indicate that the performance of the original INSIM algorithm was negatively affected not only by the use of an interior point optimizer but also by the less reliable incorporation of the physics of flow and transport than is incorporated in INSIM-FT. As in Example 1, (Tables 2.4 and 2.5), INSIM-FT gives a better data match than the other methods. In Example



(a) Prior (b) Posterior

Figure 2.20: Oil-water relative permeability curves obtained with INSIM-FT, channelized reservoir. The red solid lines represent the true relative permeability curves and blue lines represent the estimated relative permeabilities.

2, the prediction errors obtained using INSIM-FT for history-matching are significantly less than those generated with the CRM history-matched model.

Table 2.7: Well operating schedule for the Eclipse simulation model.

Well Name	Well Control	Values
I1-I4	Rate Control, RB/day	1,000
P1	BHP Control, psi	500
P2-P4	BHP Control, psi	2,000
P5	BHP Control, psi	1,000
P6	BHP Control, psi	2,500
P7	BHP Control, psi	1,000
P8	BHP Control, psi	2,000
P9	BHP Control, psi	1,000

Fig. 2.23a-2.23d, respectively, show the interwell connectivities obtained from FrontSim, CRM, INSIM and INSIM-FT respectively. Compared with the connectivity generated with CRM (Fig. 2.23b) and INSIM (Fig. 2.23c), the connectivity calculated from the INSIM-FT history-matched results in Fig. 2.23d is in better agreement with the FrontSim connectivity (Fig. 2.23a). Except for injector I4, the INSIM-FT results (Fig. 2.23d) agree well with the connectivity map generated with FrontSim based on the true Eclipse reservoir simulation model. For this example the connectivity results from INSIM match those from FrontSim better than those obtained from INSIM (ES-MDA) and because of this the INSIM (ES-MDA)

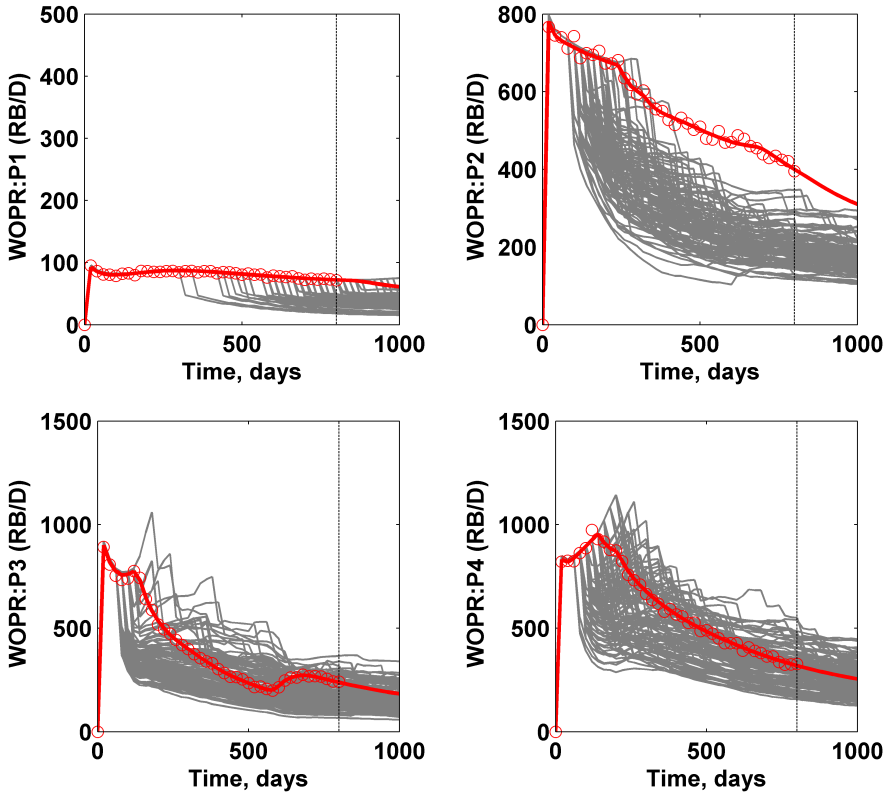


Figure 2.21: Prior oil production rate; Example 2, channelized reservoir. Red circles: observed data; red lines: true data of oil production rate; gray lines: prior responses; vertical black dashed lines: a separator for matching and prediction.

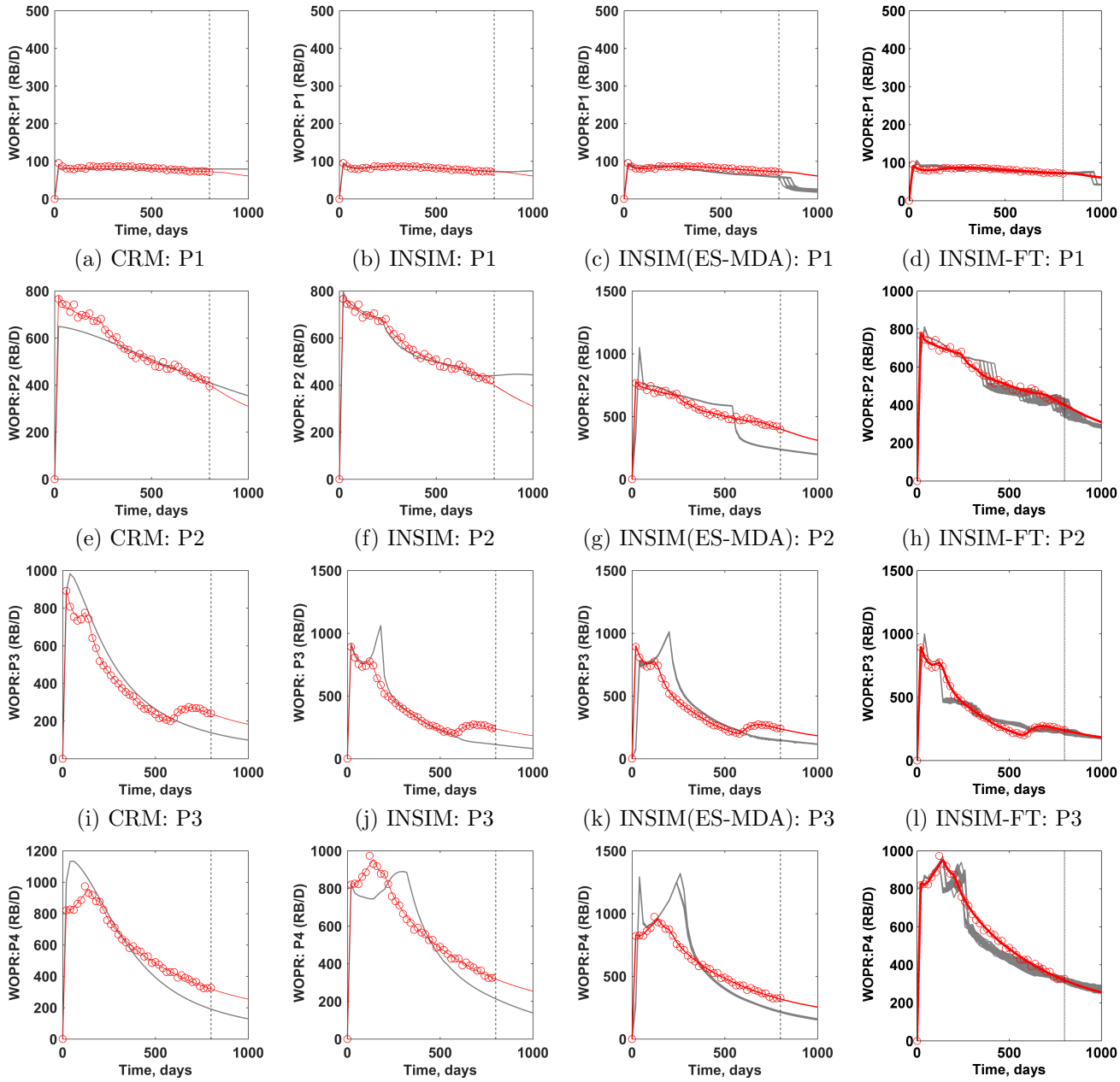
connectivity results are not shown.

2.4.3 Example 3: Field Example

This example pertains to a real reservoir with strong bottom water drive. The reservoir has 13 producing wells and no injection wells. For the purposes of history matching, we assume the aquifer is at constant pressure and represent the aquifer by a single node with the pressure at this node fixed equal to the constant pressure (p^0) of the aquifer, which is

Table 2.8: Comparison of data mismatch for historical period (800 days) and prediction period among CRM, INSIM, INSIM-FT and INSIM (ES-MDA); Example 2, channelized reservoir.

	CRM	INSIM	INSIM-FT	INSIM (ES-MDA)
Historical Period	23.28	313	10.11	62.1
Prediction Period	115.87	914	37.14	185



(a) CRM: P1 (b) INSIM: P1 (c) INSIM(ES-MDA): P1 (d) INSIM-FT: P1

(e) CRM: P2 (f) INSIM: P2 (g) INSIM(ES-MDA): P2 (h) INSIM-FT: P2

(i) CRM: P3 (j) INSIM: P3 (k) INSIM(ES-MDA): P3 (l) INSIM-FT: P3

(m) CRM: P4 (n) INSIM: P4 (o) INSIM(ES-MDA): P4 (p) INSIM-FT: P4

Figure 2.22: Comparison of oil production rates calculated with the CRM, INSIM, INSIM(ES-MDA) and INSIM-FT history-matched models for four wells; Example 2, channelized reservoir. Red circles: observed data; red lines: true data of oil production rate; gray curve: posterior oil production rate.

a procedure suggested by Zhao et al. (2016) for the original INSIM model. The value of p^0 is set equal to the value of the initial reservoir pressure at datum and does not change with time. The reservoir connection map is shown in Fig. 2.24. Note that there are 26 imaginary

Table 2.9: Comparison of data mismatch for historical period (400 days) and prediction period among CRM, INSIM, INSIM-FT and INSIM (ES-MDA) ; Example 2, channelized reservoir.

	CRM	INSIM	INSIM-FT	INSIM (ES-MDA)
Historical Period	13.8	264.9	7.6	70.5
Prediction Period	137.4	2008.4	71.7	98.4

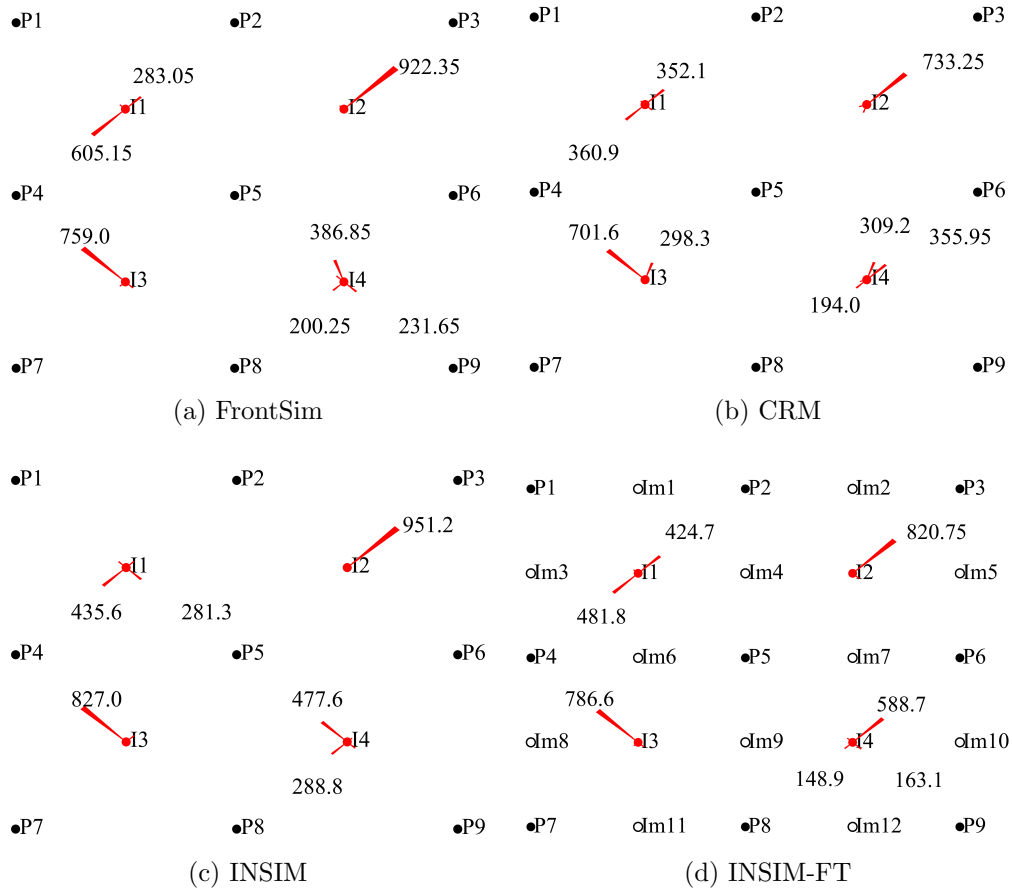


Figure 2.23: Interwell connectivity obtained from FrontSim, CRM, INSIM and INSIM-FT for the channelized reservoir. The length of a narrow red triangular denotes the magnitude of total interwell flow rate between a corresponding injector-producer pair and the direction of the triangular indicates which producer belongs to this injector-producer well pair.

wells and these imaginary wells as well as the actual wells are all connected to the single aquifer node. There are no water injection wells in the reservoir, but one might guess that CRM could be applied to this problem by treating the aquifer node as a water injection well. However, CRM requires that the total water injection rates are known and we have no knowledge of the flow rate from the aquifer to the reservoir. Please note that CRM cannot

be applied in this example since we don't know the "injection rate" for the aquifer node.

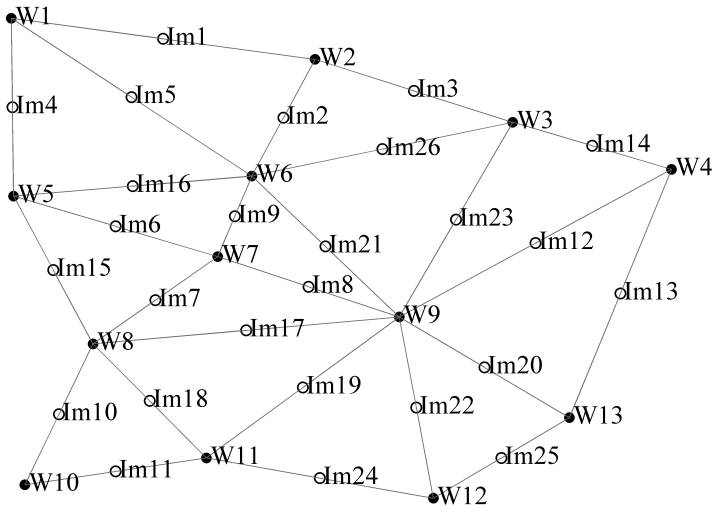


Figure 2.24: Connection map of the field example.

For our implementation of INSIM-FT, the spatial coordinates of the virtual well are computed as the average of the corresponding coordinates of the real wells as given by

$$\begin{aligned} x_v &= \frac{1}{n_w}\sum_{i}^{n_w} x_i; \\ y_v &= \frac{1}{n_w}\sum_{i}^{n_w} y_i, \end{aligned} \tag{2.38}$$

where x_v and y_v, respectively, represent the x and y coordinates of the virtual well node and n_w is the number of real wells. If well i is connected to this virtual node, the connection length is given by

$$L_{i,v} = \sqrt{(x_v - x_i)^2 + (y_v - y_i)^2}, \tag{2.39}$$

where x_i and y_i, respectively, represent x and y coordinates of well i. The virtual aquifer well node functions the same way as an injection well node. Parameters of the connections between the virtual well node and the real or imaginary wells are added to the set of parameters to be estimated by history matching. By tuning the pore volumes and transmissibilities of these connections, the water breakthrough time at a producing well can be adjusted and is not determined by the distance between the well in the basic connection map to the virtual

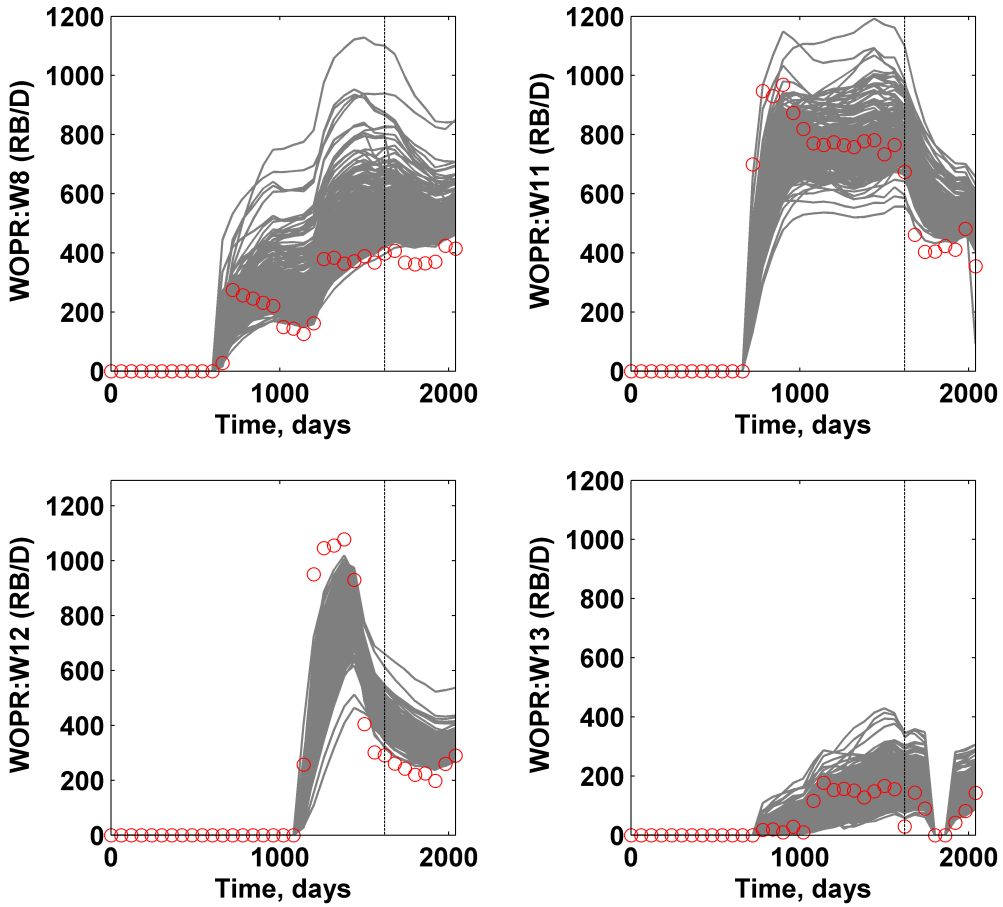
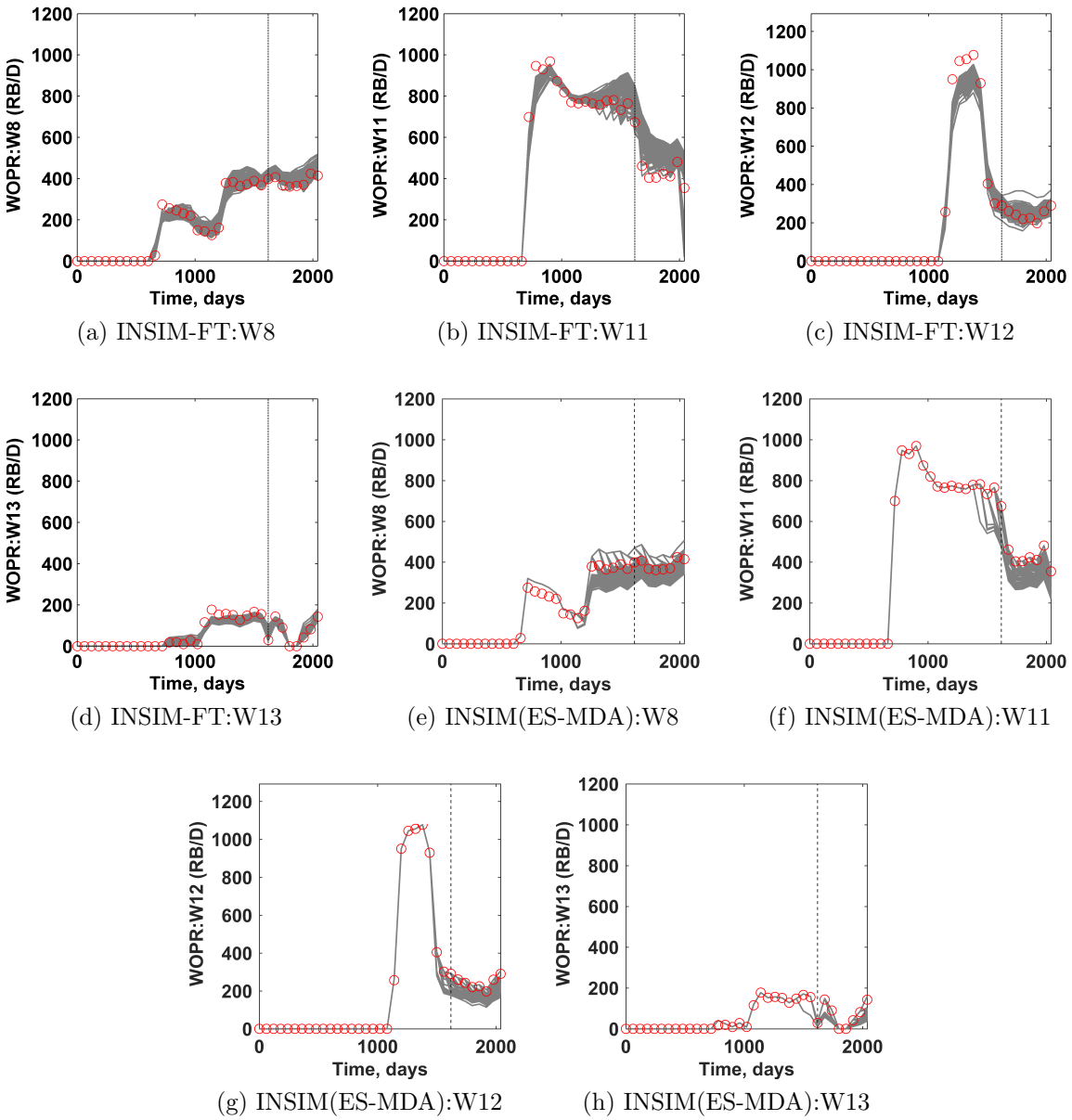


Figure 2.25: Prior oil production rates of the field example. Red circles: observed oil production rates; gray lines: prior responses of the oil production rates obtained from the prior INSIM-FT models; vertical black dashed lines: separator for matching and predictions.

well node which represents the aquifer. When history matching the field data, we assume that all connections share a global set of relative permeabilities and based on available data, the true values of parameters of the power law relative permeability curves are estimated as $a_{\text{true}} = 0.14$, $n_{w,\text{true}} = 3.0$, $n_{o,\text{true}} = 4.0$, $S_{iw,\text{true}} = 0.15$ and $S_{or,\text{true}} = 0.33$. In history matching, the values of S_{iw} and S_{or} are fixed at the estimated true values and the other three relative permeability parameters are included in the set of model parameters used for history matching. In generating initial realizations of these parameters, it is assumed that they follow the normal distributions given by $a \sim \mathcal{N}(\bar{a}, 0.05^2)$, $n_w \sim \mathcal{N}(\bar{n}_w, 0.2^2)$ and $n_o \sim \mathcal{N}(\bar{n}_o, 0.2^2)$, where the mean values, are different than the estimated true values to test the robustness of INSIM-FT and are given by $\bar{a} = 0.16$, $\bar{n}_w = 2.5$ and $\bar{n}_o = 3.3$.

The initial guess of the average reservoir absolute permeability is set equal to 100 md, which is an estimate given by field engineers. The prior ensemble of transmissibilities and pore volumes of different connections are generated in the same way as for the two synthetic examples. It should be noted that the total pore volume of the reservoir at time zero, $V^0_{p,\mathrm{tot}}$, is also tuned as a history-matching parameter for this example. The prior value of $V^0_{p,\mathrm{tot}}$, which is denoted by $\overline{V}^0_{p,\mathrm{tot}}$, is an estimate of the initial total pore volume supplied by the operator. The prior ensemble of $V^0_{p,\mathrm{tot}}$ is generated from the normal distribution $\mathcal{N}(\overline{V}^0_{p,\mathrm{tot}}, (0.2\overline{V}^0_{p,\mathrm{tot}})^2)$. The lower bound of $V^0_{p,\mathrm{tot}}$, $V^0_{p,\mathrm{tot,low}}$, is set to the half of the prior value of $V^0_{p,\mathrm{tot}}$ and the upper bound of $V^0_{p,\mathrm{tot}}$, $V^0_{p,\mathrm{tot,up}}$, is set to the 150% of the prior value of $V^0_{p,\mathrm{tot}}$. The ensemble size for ES-MDA is 300. The first 800 days of production data are history-matched and the following 200 days are used for predictions in order to ascertain whether the history-matched INSIM-FT model is capable of providing reasonable predictions of future reservoir performance. All the producing wells are operated under the specified total liquid production rates that vary with time. The observed oil production rates are the observed data to be history matched.

Representative oil production rates obtained from the prior ensemble of realizations of INSIM-FT models are shown in Fig. 2.25 whereas Figs. 2.26a-2.26d shows both the history-matching results and subsequent predictions calculated with INSIM-FT using the history-matched ensemble of models. The prior and posterior relative permeabilities are shown in Figs. 2.27a and b. The results suggest that the posterior relative permeabilities bound the true ones and are consistent with the production data. Though not shown here, the history matched values of $V^0_{p,\mathrm{tot}}$'s also bound $\overline{V}^0_{p,\mathrm{tot}}$ with a narrower uncertainty range compared with that of the prior ensemble of $V^0_{p,\mathrm{tot}}$'s. Overall, the history matched results are remarkably good. For comparison purposes, the INSIM (ES-MDA) history matched oil rates at four wells are also shown in Figs 2.26e-2.26h. A careful examination of the results indicates that better matches and predictions are obtained with INSIM-FT than with INSIM (ES-MDA); this point can be seen more clearly by examining the history matching and prediction errors given in Table 2.10 which shows the history-matching and predictions errors for INSIM-FT are less than one-half the corresponding errors obtained with INSIM (ES-MDA).



(a) INSIM-FT:W8 (b) INSIM-FT:W11 (c) INSIM-FT:W12

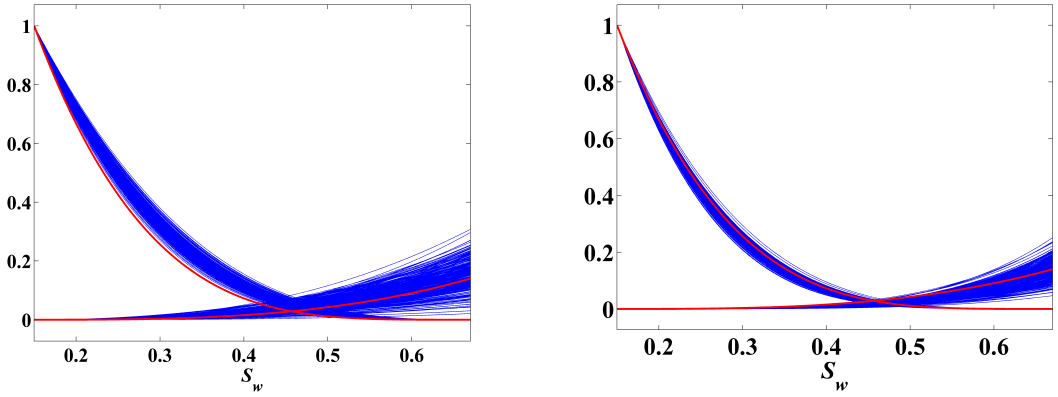
(d) INSIM-FT:W13 (e) INSIM(ES-MDA):W8 (f) INSIM(ES-MDA):W11

(g) INSIM(ES-MDA):W12 (h) INSIM(ES-MDA):W13

Figure 2.26: History-matched oil production rates of the field example obtained with INSIM(ES-MDA) and INSIM-FT. Red circles: observed oil rates; gray lines: the oil rates estimated with the history-matched INSIM-FT or INSIM(ES-MDA) models; vertical black dashed lines: separator for matching and predictions.

Table 2.10: The comparison of data mismatch for historical period and prediction period between INSIM (ES-MDA) and INSIM-FT; field example.

	INSIM	INSIM-FT
Historical Period	115.5	41.2
Prediction Period	213.8	96.6



(a) Prior (b) Posterior

Figure 2.27: Estimated relative permeabilities obtained with INSIM-FT, field example. The red lines denote the true relative permeability curves and the blue lines are the estimates of the relative permeability curves.

CHAPTER 3

WATERFLOODING OPTIMIZATION WITH THE INSIM-FT DATA-DRIVEN MODEL

In Chapter 2, we developed a physics-based data-driven model referred to as INSIM-FT and showed that it can be used for history matching and future reservoir performance predictions even when no prior geological model is available. The model requires no prior knowledge of petrophysical properties. In this chapter, we explore the possibility of using INSIM-FT in place of a reservoir simulation model when estimating the well controls that optimize water flooding performance where we use the net present value (NPV) of life-cycle production as our cost (objective) function. The well controls are either the flowing bottom-hole pressure (BHP) or total liquid rates at injectors and producers on the time intervals which represent the prescribed control steps. The optimal well controls that maximize NPV are estimated with an ensemble-based optimization algorithm using the history-matched INSIM-FT model as the forward model.

To develop the INSIM-FT model from observed production data, the relative permeability parameters, as well as the connection-based parameters are estimated using the ensemble smoother with multiple data assimilation (ES-MDA) (Emerick and Reynolds, 2012, 2013a,b). Here, we show that if bottom-hole pressure data are available, we can estimate the values of well indices via the history matching process. This procedure to estimate well indices is critical in order to use the history-matched INSIM-FT model for production optimization for the case where the well controls are bottom-hole pressures and thus is a vital contribution of the current work.

In each example considered, the production data that we history match is generated from an Eclipse 100 reservoir simulation model which is used to represent the true reservoir

model. To illustrate the potential of using INSIM-FT for life-cycle production optimization, the optimal well controls obtained by optimization using INSIM-FT as the forward model are input to the true reservoir model (reservoir simulation model) to compute future production and the resulting life-cycle NPV of production. Also, we compare the optimal NPV obtained using INSIM-FT as the forward model with the estimate of the optimal NPV obtained using the correct full-scale reservoir simulation model when performing waterflooding optimization. The optimization method applied to maximize NPV is the EnOpt algorithm of Lorentzen et al. (2006); Chen et al. (2009) based on the very minor modification introduced by Do and Reynolds (2013).

3.1 Waterflooding Optimization Problem

Waterflooding is an important secondary recovery method which maintains reservoir pressure and increases oil recovery. Life-cycle waterflooding optimization aims to find well controls (operating well pressures or rates at specified time intervals (control steps)) that maximize some cost function, e.g., net-present-value of production or cumulative oil production over the productive life of the reservoir. In this work, the cost function is defined as the net-present-value (NPV) of production which is given by

$$J(u) = \sum_{n=1}^{N_t} \left\{ \frac{\Delta t_n}{(1+b)^{\frac{t_n}{365}}} \Big[\sum_{j=1}^{P} (r_o \cdot q_{o,j}^n - c_w \cdot q_{w,j}^n) - \sum_{j=1}^{I} (c_{wi} \cdot q_{wi,j}^n) \Big] \right\}, \tag{3.1}$$

where r_o (\$/RB) is the oil revenue; c_w (\$/RB) is the cost of disposing of produced water; c_{wi}(\$/RB) is the water injection cost and b is the annual discount rate. N_t is the total number of simulation time steps; t_n is the nth time level and Δt_n is the length of the nth time step. P and I, respectively, denote the number of producers and injectors; $q_{o,j}^n$ (RB/D) and $q_{w,j}^n$ (RB/D) denote the average oil and water production rate at the jth producer at the nth time step; $q_{wi,j}^n$ (RB/D) denotes the average water injection rate at the jth injector at the nth time step. It is important to note that in Eq. 3.1, c_w should not be confused with water isothermal compressibility.

For the problem considered, we perform production optimization after the reservoir

has been produced for a sufficiently long time to collect and history-match production data in order to determine a model using the INSIM-FT methodology and then this model (INSIM-FT) is used as the forward model when estimating the optimal well controls that maximize NPV over the remaining life of the reservoir; throughout this chapter, this optimization step is referred to as life-cycle production optimization. Here, the time-period for production optimization is divided into N_c consecutive time intervals of equal length, where each time interval is referred to as a control step. On each control step, one needs to specify the operating condition for each well, where the operating condition (well control) may be a bottom-hole pressure (BHP) or a total or phase rate. During a specific control step (time interval), each well is operated with a constant value of the well control but the value of the well control may differ from well to well as well as from control step to control step. We let the vector u denote the vector of all control variables, where, for convenience in presentation, u is ordered as

$$u = \begin{bmatrix} u^1 \\ u^2 \\ . \\ . \\ u^{N_w} \end{bmatrix}, \tag{3.2}$$

where N_w is the number of wells and

$$u^\ell = \begin{bmatrix} u_1^\ell \\ u_2^\ell \\ . \\ . \\ u_{N_c}^\ell \end{bmatrix}, \tag{3.3}$$

where u_j^ℓ denotes the well controls for well ℓ at the jth control step for $j = 1, 2, \ldots, N_c$ and

$\ell = 1, 2, \ldots, N_w$. The dimension of u is denoted by N_u and given by

$$N_u = N_c \cdot N_w.$$

Here, we assume that these variables are constrained by simple bounds, so that the waterflooding optimization problem is expressed as

$$\underset{u \in \mathbb{R}^{N_u}}{\text{maximize}}\ J(u), \tag{3.4}$$

subject to

$$u_k^{\ell,\text{low}} \leq u_k^{\ell} \leq u_k^{\ell,\text{up}}, \tag{3.5}$$

where $u_k^{\ell,\text{low}}$ and $u_k^{\ell,\text{up}}$ are the lower and upper bounds of kth control variable for well ℓ. To attempt to achieve better scaling, these control variables are normalized so they are replaced by

$$x_k^{\ell} = \frac{u_k^{\ell} - u_k^{\ell,\text{low}}}{u_k^{\ell,\text{up}} - u_k^{\ell,\text{low}}}. \tag{3.6}$$

Note that for the normalized variables, the lower and upper bounds, respectively, are 0 and 1, for $k = 1, 2, \ldots, N_c$ and $\ell = 1, 2, \ldots, N_w$.

3.2 Average History-Matched INSIM-FT Model

In Chapter 2, the ensemble smoother with multiple data assimilation (ES-MDA) (Emerick and Reynolds, 2012, 2013a,b) was adopted to history match the production history of a reservoir to obtain an ensemble of posterior models. If we assume all the well connections share the same set of relative permeability functions, the vector containing all the model parameters is given by

$$m = [T_{i,j}^0\text{'s}, V_{p,i,j}^0\text{'s}, a, n_w, n_o]. \tag{3.7}$$

In history matching with INSIM-FT, the total liquid production rate as a function of time at producing wells and the total water injection rate at injection wells as a function of

time are specified in order to define the source/sink terms in Eq. A-18 and water saturation are solved along the connections using front tracking (Holden et al., 1988; Lie and Juanes, 2005; Guo et al., 2017c). After history matching, we use the INSIM-FT model based on the average of the posterior realizations of m to perform life-cycle production optimization. The average model, m_{avg}, is defined as

$$m_{\text{avg}} = \frac{1}{N_e} \sum_{l=1}^{N_e} m_l, \tag{3.8}$$

where the subscript i denotes the ith posterior model and N_e is the total number of posterior models obtained with INSIM-FT.

3.3 Estimation of Optimal Well Controls

Normally, for well-control production optimization, well controls can be either bottom-hole pressure or rate. However, because INSIM-FT assumes the total injection or production rate is specified at each well, to use pressures as control variables requires us to replace the source/sink term in Eq. A-18 with one involving pressure. Specifically, we replace $q^n_{t,i}$ in Eq. A-18 with the right-hand side of the following equation:

$$q^n_{t,i} = -\text{WI}_i \lambda^{n-1}_{t,i} (p^n_i - p^n_{wf,i}) \tag{3.9}$$

to obtain

$$\frac{c^{n-1}_{t,i} V^{n-1}_{p,i}}{\Delta t_n} \left(p^n_i - p^{n-1}_i\right) = \sum_{j=1}^{n_{c,i}} T^{n-1}_{i,j} (p^n_j - p^n_i) - \text{WI}_i \lambda^{n-1}_{t,i} (p^n_i - p^n_{wf,i}), \tag{3.10}$$

where the bottom-hole pressure $p_{wf,i}$ can now be specified as the well control for well i at each time level. Unfortunately, we cannot use Peacemans equation (Peaceman, 1978, 1983) to calculate the well indices because there are no permeabilities or gridblock dimensions explicitly involved in the INSIM-FT model. Thus, we develop below a procedure for estimating the well indices, WI's in Eq. 3.10 in order to perform production optimization. In Eq. 3.9, $q^n_{t,i}$ is the total rate that was specified during the history-matching process. Assuming wellbore

Table 3.1: Mean values and standard deviations of well indices for Example 1.

	P1	P2	P3	P4
Mean, 10^3 md ·ft	3.42	1.36	3.67	1.30
Standard Deviation, 10^3 md ·ft	0.239	0.176	0.198	0.130

pressure measurements are available, we can solve Eq. 3.9 directly for WI's using the p_i^n values calculated with the history-matched model, the measured $p_{wf,i}^n$ values and the $\lambda_{t,i}$ values computed from the saturations obtained from the history match run. The primary difficulty with this procedure is that due to the approximate nature of INSIM-FT, a WI_i calculated from Eq. 3.9 varies slightly with n (Guo et al., 2018b), whereas, well indices are by definition independent of time. For example, for the first example we test later, the mean values and standard deviations of well indices for P1 to P4 obtained from the history-match run by Eq. 3.9 are shown in Table 3.1, where the standard deviations are fairly small compared to the mean values. Therefore, instead of solving Eq. 3.9, for the ith well index, we use an average over time, i.e., we compute the ith well index as

$$\mathrm{WI}_i = \frac{1}{n_t} \sum_{k=1}^{n_t} \frac{-q_{t,i}^k}{(p_i^k - p_{wf,i}^k)\lambda_{t,i}^k}, \tag{3.11}$$

where n_t is the number of INSIMS-FT steps in the history-matching period.

When we perform history matching with ES-MDA, we obtain multiple models, i.e., multiple realizations of the parameters defining an INSIM-FT model. As our objective here is simply to demonstrate that the INSIM-FT model can be used to solve the optimal well control problem, we first evaluate Eq. 3.11 based on running INSIM-FT with m_{avg} defined in Eq. 3.8, then we perform long-term production optimization (water flooding optimization) based on m_{avg}. With the WI_i's computed from Eq. 3.11, we can now use pressure controls. At each iteration of the optimization algorithm, we obtain a new estimate of all well pressure controls, i.e., the values of the $p_{wf,i}$'s used in Eq. 3.10 to run INSIM-FT to calculate pressures, p_i^n's. After solving the INSIM-FT model for pressure and water saturation as presented in Chapter 2, the total flow rate at injectors and producers can be calculated by

$$q_{t,i}^n = \mathrm{WI}_i \cdot \lambda_{t,i}^{n-1}(S_{w,i}^{n-1}) \cdot (p_i^n - p_{wf,i}^n). \tag{3.12}$$

Then, the oil production rates at time level n is computed by

$$q_{o,i}^n = q_{t,i}^n \cdot (1 - f_w(S_{w,i}^n)). \tag{3.13}$$

Because we neglect gravity and capillarity effects, f_w is given by

$$f_w(S_w) = \frac{\frac{k_{rw}(S_w)}{\mu_w}}{\frac{k_{rw}(S_w)}{\mu_w} + \frac{k_{ro}(S_w)}{\mu_o}}, \tag{3.14}$$

where μ_w and μ_o are constant and known, and $k_{rw}(S_w)$ and $k_{ro}(S_w)$, respectively, are given by Eqs. 2.17 and 2.18. The water production rate is computed as

$$q_{w,i}^n = q_{t,i}^n \cdot f_w(S_{w,i}^n). \tag{3.15}$$

Given these rates, the NPV value of production (Eq. 3.1) can be calculated for the current estimate of the optimal well controls. For the two synthetic examples considered later, we use rate controls at water injection wells and pressure controls at producing wells.

3.4 Optimization Procedure

3.4.1 Ensemble-Based Optimization (EnOpt).

EnOpt (Chen et al., 2009) based on the implementation of Do and Reynolds (2013) is used to estimate the optimal controls that maximize the NPV of life-cycle production. Optimization is done in terms of the normalized control variables defined in Eq. 3.6 and then, given the value of the normalized control variable x_j^ℓ, Eq. 3.1 is solved for the corresponding value of u_j^ℓ. Throughout, x represents the vector of normalized variables corresponding to the control variable u defined in Eq. 3.2. With EnOpt, at iteration k of the steepest ascent

algorithm, the search direction is based on a smoothed stochastic gradient d_k of the NPV objective function J defined in Eq. 3.1. This search direction is normalized by the infinity norm so the steepest ascent algorithm becomes

$$x^{k+1} = x^k + \alpha_k \left[\frac{d_k}{\|d_k\|_\infty} \right], \tag{3.16}$$

for $k = 0, 1, \ldots,$ until convergence, where α_k is the step size and is determined by a line search with a back tracking strategy. Since in Eq. 3.6, the values of each normalized control variables must be in the interval $[0, 1]$, the search direction d_k is also normalized by its infinity norm so that, if $\alpha_k \leq 1$, the maximum change difference in any component of $x^{k+1} - x^k$ is less than or equal to unity. To compute d_k, an ensemble of normalized control variables needs to be randomly generated. To avoid abrupt changes in a well's control from one time step to another, during the generation, we often impose temporal correlation on controls at each well by introducing a covariance matrix, where C_x^ℓ denotes the covariance matrix for well ℓ. Denoting $C_{i,j}^\ell$ as the (i, j) entry of C_x^ℓ, $C_{i,j}^\ell$ is defined here by the spherical covariance function given by

$$C_{i,j}^\ell = \begin{cases} \sigma^2 \left[1 - \frac{3}{2} \frac{|i-j|}{N_s} + \frac{1}{2} \left(\frac{|i-j|}{N_s} \right)^3 \right] & |i - j| \leq N_s \\ 0 & |i - j| > N_s \end{cases}, \tag{3.17}$$

where i and j, respectively, refer to the ith and jth control step for well ℓ; N_s is the number of control steps over which we wish the control to be correlated; σ is the standard deviation, which is set to 0.01 for our application. The overall covariance matrix is the block diagonal matrix given by

$$C_x = \begin{bmatrix} C_x^1 & 0 & . & 0 \\ 0 & C_x^2 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & C_x^{N_w} \end{bmatrix}. \tag{3.18}$$

By sampling the Gaussian random distribution $\mathcal{N}(x^k, C_x)$ N_p times, we obtain the smoothed ensemble of normalized well controls, denoted by $\hat{x}^j$, $j = 1, 2, \ldots, N_p$. If a sampled $\hat{x}^j$ value is not in [0,1], then we simply implement truncation to enforce bound constraints. With truncation, whenever a component of $\hat{x}^j$ is greater than its upper bound 1, we set that component equal to 1, and whenever a component is less than its lower bound 0, we set that component equal to 0. In this work, we apply a slightly modified version of EnOpt (Chen et al., 2009) (see Do and Reynolds (2013)) to compute

$$d_k = \frac{C_x}{N_p} \sum_{j=1}^{N_p} (\hat{x}^j - x^k)[J(\hat{x}^j) - J(x^k)], \tag{3.19}$$

where J is the objective function given by Eq. 3.1.

3.4.2 Validation.

A data-driven model does not have the power to exactly reproduce the predictions of the full-scale reservoir simulator. What we really care about is whether the optimal well controls obtained by INSIM-FT can also give a significant improvement in the NPV of production for the true model. Since all the synthetic observed data are generated by running an Eclipse 100 truth model and adding noise, the Eclipse simulation model represents the true model. Intuitively, we have two ways to validate the INSIM-FT production optimization results: (1) input the INSIM-FT optimal well controls into the Eclipse schedule and compute the corresponding NPV; (2) perform production optimization independently on the true Eclipse model with Eclipse as the forward model to obtain the optimized NPV.

3.4.3 Complete Workflow.

The complete workflow for production optimization step is summarized below.

1. Specify the maximum number of total iterations, $N_{maxiter}$, allowed in steepest ascent and the maximum number of simulation runs, N_{maxsim}, allowed for optimization.

2. Choose the average history matched INSIM-FT model obtained using ES-MDA for production optimization. Run INSIM-FT with the total rates specified as the historical rates for this model and obtain the well indices from Eq. 3.11.

3. Start production optimization and set $k = 0$, where k is the iteration index for steepest ascent (Eq. 3.16). In well production optimization example presented here, well controls during the optimization period are set to BHP controls for producers and rate controls for injectors. Use the observed BHP data and total injection and production rate data at the end of production history to estimate the initial guesses of controls for all control variables and generate the initial normalized control vector x^0.

   For $k = 0, 1, \ldots, N_{maxiter}$,

4. Generate N_p perturbations of the normalized control vector, x^k, by sampling from the Gaussian distribution $\mathcal{N}(x^k, C_x)$; the perturbations are given by

$$\hat{x}^j = x^k + C_x^{1/2} Z_j, \;\; j = 1, 2, \ldots, N_p,$$

   where

$$Z_j \sim \mathcal{N}(0, \mathrm{I}).$$

   If a component of $\hat{x}^j$ is outside the interval of [0,1], apply truncation to ensure all components of $\hat{x}^j$ satisfy the bound constraints.

5. Run INSIM-FT for each set of well controls, $\hat{u}^j = \hat{x}^j \left[u^{\mathrm{up}} - u^{\mathrm{low}}\right] + u^{\mathrm{low}}$, $j = 1, 2, \ldots, N_p$ to obtain corresponding NPV values.

6. Compute the EnOpt gradient given by Eq. 3.19.

7. Do line search with a back tracking to obtain the updated control vector x^{k+1}. The initial stepsize is set to 0.1. If the line search can not find a control vector which increases the value of NPV in the maximum allowable number of stepsize cuts, then

we set x^{k+1} to the set of controls obtained during the line search that gives the highest NPV. In the examples we consider, the maximum number of stepsize cuts is set to 5.

8. If the number of maximum simulation runs N_{maxsim} or maximum iteration $N_{maxiter}$ are reached, or both of the following conditions are satisfied,

$$\frac{|J(x^{k+1}) - J(x^k)|}{\max[J(x^k), 1.0]} \leq \varepsilon_J, \tag{3.20}$$

$$\frac{\|x^{k+1} - x^k\|}{\max[\|x^k\|_2, 1.0]} \leq \varepsilon_x, \tag{3.21}$$

then terminate iteration. In the examples, we use $\varepsilon_J = 10^{-4}$ and $\varepsilon_x = 10^{-3}$.

End For

9. Record the values of u for x^k that gives the highest NPV as the optimal well controls.

3.5 Examples

We use the Eclipse 100 simulator to validate our INSIM-FT production optimization algorithm. The three examples that are tested for history matching with INSIM-FT in Chapter 2 now are used to test the performance of production optimization with INSIM-FT. Specifically, EnOpt is performed with INSIM-FT using the average of the history-matched models in INSIM-FT to obtain the optimal well controls for water injectors and producers.

3.5.1 A Homogeneous Reservoir with a Sealing Fault

The reservoir with a sealing fault is shown in Fig. 2.5 in Chapter 2. The Eclipse true model is based on a $33 \times 33 \times 1$ grid with grid dimensions being $\Delta x = \Delta y = 100$ ft and $\Delta z = 50$ ft and the fault is represented by the white inactive cells. The reservoir has a homogeneous permeability field of 500 mD and porosity field of 0.2 with nine total wells, five water injectors and four producing wells. A continuous zigzag sealing fault divides this reservoir into two parts. During the historical period when production data are obtained, 5

injectors are all operated at a constant rate of 1,000 RB/day and 4 producers are all operated at a constant bottom-hole pressure of 2,000 psia. The other reservoir properties are specified in Table 2.3. The true data, including the total production rate and oil production rate, are obtained by running Eclipse 100 for 1,000 days with the true grid-based parameters specified. The observed oil production rates are obtained by adding uncorrelated Gaussian noise to the true oil production rates obtained from Eclipse where the standard deviation of each measurement error is set equal to 2% of the true data.

History matching. Since INSIM-FT for history matching is run based on rate controls, to history match the observed data, the true water injection rate and the true total production rate are specified in INSIM-FT and the observed oil production rate data from Eclipse are history matched to obtain the vector of model parameter, m defined in Eq. 3.7, for INSIM-FT. The connection map is shown in Fig. 2.14. Im1-Im4 are four imaginary wells added to increase more flow channels between injectors and producers; see Chapter 2. Here, all the connections share the same set of relative permeability curves (see Eqs. 2.17-2.19).

History matching is performed using ES-MDA starting with 100 sets of prior model parameters (transmissibilities, pore volumes and relative permeability parameters). The history matching period lasts for 1,000 days and the detailed procedure is shown in Chapter 2.

Production Optimization. Production optimization which aims to maximize the objective function defined in Eq. 3.1, is performed following the 1,000 days of historical production based on a total reservoir life of 2,000 days, i.e, the optimization time interval is equal to 1,000 days.

The well indices for pressure control are computed by Eq. 3.11 from a single forward run of INSIM-FT with m_{avg}. The control variables of injectors are injection rates at reservoir conditions with an upper bound of 3,000 RB/day and a lower bound of 0 RB/day; the control variables of producers are bottom-hole pressure with an upper bound of 3,000 psia and a lower bound of 200 psia. The 1,000 days optimization period is divided into 10 control steps and hence each control step lasts 100 days and therefore the control vector u has $9 \times 10 = 90$

elements. The initial guesses for well controls are given by the well controls at the end of matching period, which are 1,000 RB/day for injection rates of all water injectors and 2,000 psia for the BHP of all producers.

An optimal control strategy is obtained by maximizing the NPV as described in Eq. 3.1, with $r_o = 80$ USD/RB, $c_w = 5$ USD/RB, $c_{wi} = 2$ USD/RB. The discount rate $b = 0.1$. For EnOpt, the ensemble size for gradient estimation is $N_p = 10$. The maximum number of stepsize cuts for line search with backtracking is 5; the initial search step for steepest ascent is 0.1; the time-correlation length is 300 days; the maximum number of allowable simulation runs N_{maxsim} is equal to 1,000; the maximum number of optimization iterations, $N_{maxiter}$, is equal to 60.

To validate the optimization results obtained using the average of the INSIM-FT history-matched models, we perform independent production optimization with Eclipse 100 by using EnOpt with the same initial optimization settings as is used in optimization with INSIM-FT. The reservoir simulation model used in Eclipse to perform production optimization is the true model. The comparison of NPV versus the number of reservoir simulation runs when optimizing with INSIM-FT and optimizing with Eclipse are shown in Fig. 3.1. For this particular example, the total computational time for production optimization with Eclipse is one hour and five minutes but for INSIM-FT is only one minute. Within the same number of iterations, INSIM-FT achieves about a 45 million USD higher NPV value than the NPV obtained with the initial control settings, while Eclipse with independent optimization increases the initial NPV by 25 million USD. Due to the model difference between Eclipse and INSIM-FT, the estimated well controls differ and the difference in NPV is not surprising. Figs. 3.2 and 3.3 show the comparisons of optimal well controls for INSIM-FT and Eclipse. In Figs. 3.2 and 3.3 and similar figures presented later, each colored rectangle represents the value of the optimal well control at a given control step and a given well. As shown in Fig. 3.2, the optimal injection rates generated from INSIM-FT and Eclipse are qualitatively similar.

A better way to validate the optimized results obtained with INSIM-FT is to apply

the optimal well controls from INSIM-FT directly in the true Eclipse 100 to compute the true production and NPV that will be obtained using the optimal well controls generated based on optimization with INSIM-FT used as the forward model. The NPV value obtained by this procedure is referred to as Eclipse V1 in Table 3.2 and Eclipse V2 refers to optimization results obtained using Eclipse directly with the true reservoir model. In Table 3.2, we show both the NPV values based on the initial guess as well as the maximum NPVs of production by three procedures. In Table 3.2, it is interesting to note that using the optimal well controls estimated with the INSIM-FT model in Eclipse gives a value of NPV equal to 274.5 million USD which is only 2.4% less than the NPV generated by optimizing well controls based on the true Eclipse reservoir simulation model, even though the optimal NPV estimated using INSIM-FT is ten percent higher than the two values of NPV generated with Eclipse. Fig. 3.4 shows that by applying the optimal well controls from INSIM-FT in the true reservoir model, a considerable amount of additional oil is produced, but the cumulative oil production from Eclipse is less than that predicted by INSIM-FT. However, the cumulative oil production obtained by independent optimization of the Eclipse model and by applying the optimal well controls from INSIM-FT in Eclipse are in very good agreement. In Fig. 3.5, we compare the oil saturation distributions obtained at the end of the production life by using the optimal well controls generated with Eclipse and INSIM-FT in the true reservoir simulation model. Even though the optimal well controls generated with INSIM-FT and Eclipse are very different, the oil saturation distributions are quite similar. This is not surprising as production optimization problems with a large number of control variables often have multiple solutions.

Since one single run may not fully demonstrate that the results are reliable, we rerun the example with INSIM-FT and Eclipse with different guesses of the well controls. The results are shown in Table 3.3. We see from Table 3.3, even with two very different initial conditions for optimization, the NPVs predicted by inputting the optimal well controls obtained with INSIM-FT into the true model are still quite similar to those obtained by direct optimization with the true model.

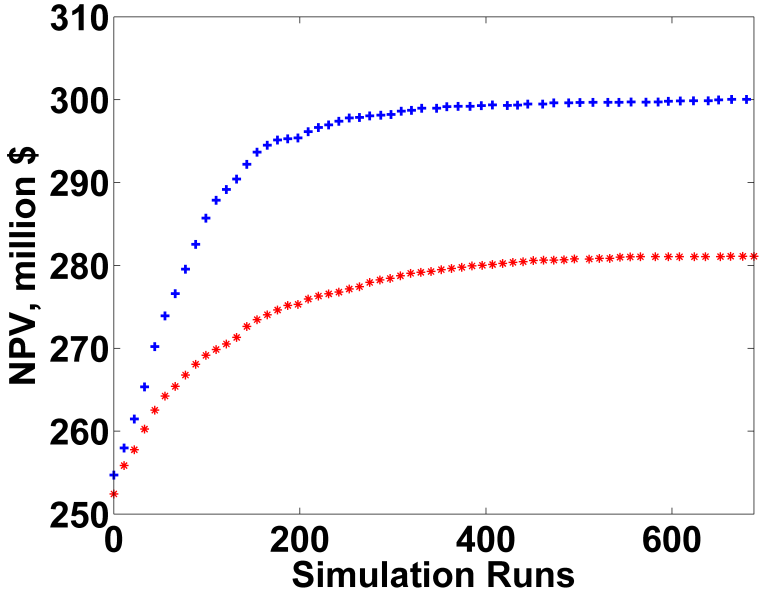
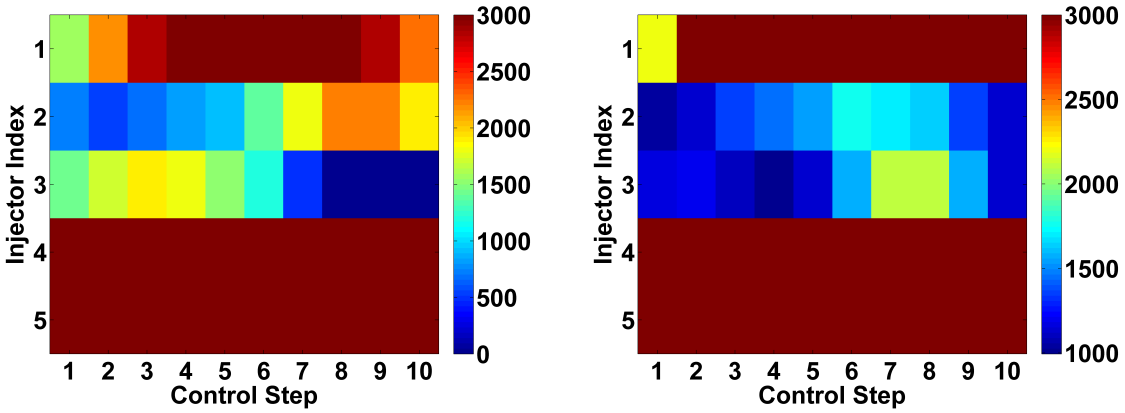


Figure 3.1: NPV versus simulation runs, fault case. Red stars denote the independent optimization with Eclipse; blue + denote the optimization with INSIM-FT.



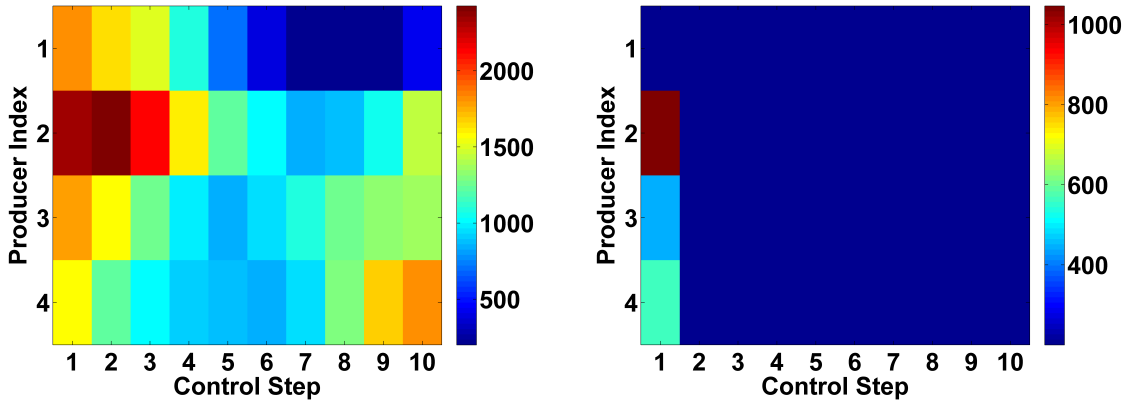
(a) Optimal well controls for injectors with INSIM-FT model

(b) Optimal well controls for injectors with Eclipse simulation model

Figure 3.2: Estimated optimal water injection rate for injectors at different control steps.

3.5.2 Channelized Reservoir

In this example, production optimization is performed for the same channelized reservoir as shown in Chapter 2. The reservoir simulation model, which represents the true reservoir is based on a 225 by 225 by 1 grid where the size of each of these gridblock is 90 ft by 90 ft by 10 ft where 10 ft is the thickness of the gridblock. The reservoir has a homogeneous porosity field with $\phi = 0.2$. The log permeability field and well locations are shown in



(a) Optimal well controls for producers obtained with INSIM-FT model

(b) Optimal well controls for producers with Eclipse simulation model

Figure 3.3: Estimated optimal BHP for producers at different control steps.

Table 3.2: NPV comparison between INSIM-FT and Eclipse validations. INSIM-FT initial and optimal, respectively, represent the NPVs generated with INSIM-FT using the initial and optimal INSIM-FT well controls. Eclipse V1 inital and optimal, respectively, represent the results obtained by running Eclipse with the same initial and optimal well controls from INSIM-FT; Eclipse V2 initial denotes the NPV generated using the initial well controls used for INSIM-FT in the true Eclipse simulation model. Eclipse V2 optimal denotes the results by doing optimization directly on the reservoir simulation model (the truth).

	INSIM-FT		Eclipse V1		Eclipse V2	
	Initial	Optimal	Initial	Optimal	Initial	Optimal
NPV, million USD	254.81	300.64	254.70	274.5	254.70	281.1

Table 3.3: Optimal NPVs obtained with INSIM-FT and Eclipse with two different initial conditions of well controls. INSIM-FT represents the optimal NPV generated with INSIM-FT; Eclipse V1 represents the NPV obtained by running Eclipse with the optimal well controls from INSIM-FT; Eclipse V2 denotes the NPV by optimizing directly on the reservoir simulation model (the truth), faulted reservoir.

	Injection Rate	Producer BHP	INSIM-FT	Eclipse V1	Eclipse V2
Case 1	900 RB/day	1,040 psi	302×10^6 USD	268×10^6 USD	278×10^6 USD
Case 2	2,100 RB/day	1,800 psi	300×10^6 USD	264×10^6 USD	276×10^6 USD

Fig. 2.18. The yellow to red areas represent high permeability channels and dark blue areas represent low permeability zones. As discussed in Chapter 2, to increase the resolution of INSIM-FT, we use an additional 12 imaginary wells, which simply represent nodes (control volumes) through which fluid can flow to a production well. The total set of connections is

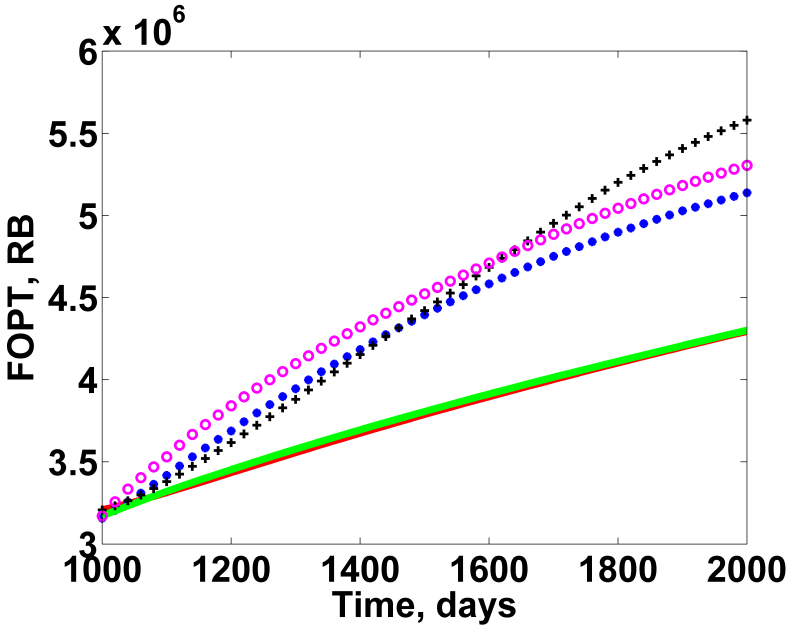
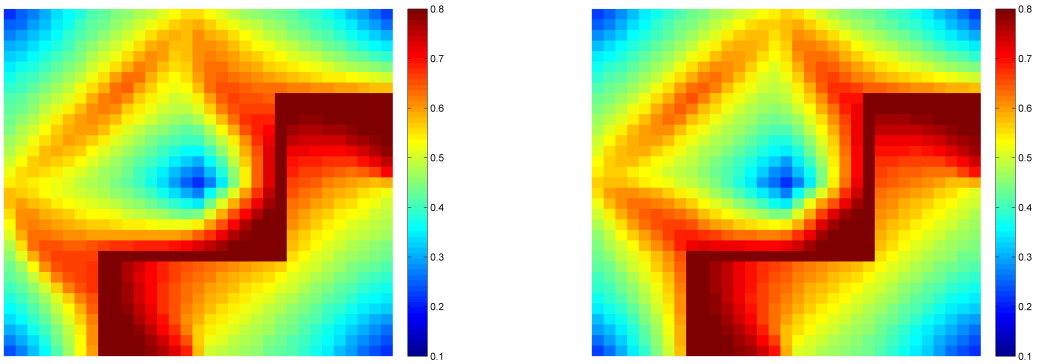


Figure 3.4: Cumulative oil production versus time, Example 1. Green curve represents cumulative oil production obtained from the initial guess of optimal well controls using INSIM-FT as the forward model; red curve represents the cumulative oil production from Eclipse with the initial guess of optimal well controls; black pluses denote the results calculated with the optimized INSIM-FT control using the INSIM-FT forward model; blue circles denote the results from Eclipse true model using the optimal controls estimated with INSIM-FT; pink circles denote the results from independent optimization by Eclipse.



(a) INSIM-FT (b) Eclipse

Figure 3.5: Oil saturation distributions obtained at the end of production life by applying the optimal well controls from INSIM-FT and Eclipse into the true faulted reservoir model.

shown in Fig 2.19, where a node labeled IM_j does not contain an actual well. The reservoir has 1,000 days of production history for history matching and the subsequent 1,000 days are used for life-cycle production optimization. The well operating conditions for historical

period are specified in Table 3.4 and other reservoir properties are shown in Table 2.6. Again the true oil production rate is obtained by running Eclipse 100 for 1,000 days and the observed oil production rate is generated by adding uncorrelated Gaussian random noise with standard deviation equal to 2% of the true data.

Table 3.4: The well operating conditions for historical period, channelized reservoir

Producer Name	Control Type	Target Value
I1	Rate, RB/day	1,000
I2	Rate, RB/day	1,000
I3	Rate, RB/day	1,000
I4	Rate, RB/day	1,000
P1	BHP, psia	500
P2	BHP, psia	2,000
P3	BHP, psia	2,000
P4	BHP, psia	2,000
P5	BHP, psia	1,000
P6	BHP, psia	2,500
P7	BHP, psia	1,000
P8	BHP, psia	2,000
P9	BHP, psia	1,000

History matching. History matching is performed using ES-MDA (Emerick and Reynolds, 2012, 2013a,b) using 8 data assimilation steps with equal inflation factors with INSIM-FT and with 100 realizations of the INSIM-FT vector of model parameters. The details can be seen in Chapter 2.

Production optimization. The production optimization performed with INSIM-FT is based on the average posterior model defined in Eq. 3.8. The objective function for this maximization problem is given by Eq. 3.1 with bound constraints to the control variables. The control variables for injectors are injection rates under reservoir conditions with an upper bound of 2,000 RB/day and a lower bound of 0 RB/day; the control variables of producers are bottom-hole pressures with an upper bound of 3,000 psi and a lower bound of 100 psi. The optimization period is divided into 10 control steps and each step lasts 100 days. Since there are 9 producers and 4 injectors, the control vector u has $13 \times 10 = 130$ elements. The initial guesses for well controls are the well controls specified at the end of the production

history. The EnOpt settings as the same as for the fault case: the number of perturbation for gradient estimation (see Eq. 3.19) is $N_p = 10$; the maximum number of stepsize cuts is 5; the initial stepsize for each steepest ascent direction is 0.1; the time-correlation length is 300 days; the maximum number of allowable simulation runs is 1,000; the maximum number of optimization iterations is 60.

As in the first example, we compare the optimization results generated with INSIM-FT to those obtained by optimization using the true Eclipse reservoir simulation model. For both optimizations, we use the same initial estimate of the optimal well controls. Fig. 3.6 shows the NPV versus the number of forward reservoir simulation runs during optimization with the INSIM-FT forward model and with Eclipse 100. The computational time required by direct optimization with Eclipse is 23.5 hours whereas INSIM-FT only requires 5 minutes to complete the optimization run. In contrast to the result for the fault case, the optimization performed with Eclipse yields a NPV value 5% higher than the NPV generated with INSIM-FT. Due to model difference between Eclipse and INSIM-FT, the difference in NPV is understandable. The comparison of optimal well controls for injectors is shown in Fig. 3.7 and for producers is shown in Fig. 3.8. Due to the model differences, the controls are quite different.

We also compute the NPV by applying the optimal well controls from INSIM-FT directly in the true reservoir simulation model to compute the corresponding true NPV, which is referred to as Eclipse V1 optimal in Table 3.5. Eclipse V2 refers to the optimal NPV obtained by optimization directly with the true Eclipse model. In Table 3.5, we show both the NPV values based on the initial well controls as well as the maximum NPV of production obtained with three procedures. From the Eclipse V1 results of Table 3.5, we see that by inputting the optimal well controls estimated with the INSIM-FT model into Eclipse, we obtain a NPV value of 234.04 million dollar which is quite close to the NPV of 241.28 million dollar generated by optimization with the true Eclipse reservoir simulation model. To further validate the optimization performance of INSIM-FT, optimization is done with two different initial conditions and the results are shown in Table 3.6. In general,

the optimal controls from INSIM-FT yield an estimate of optimal NPV that is very similar to the NPV generated by optimization directly with the "true" reservoir simulation model. Fig. 3.9 shows that by applying the optimal well controls from INSIM-FT in the true reservoir simulation model, we obtained a cumulative oil production profile, which is in good agreement to the one obtained by production optimization directly with the true Eclipse reservoir simulation model. Compared with the two Eclipse-generated results, INSIM-FT predicts less oil production using the optimal INSIM-FT well controls.

Fig. 3.10 compares the oil saturation distributions obtained at the end of the optimization period with optimal controls generated with INSIM-FT and those generated with Eclipse. As shown in Fig. 3.10, the oil saturation distributions for the two alternative methods look qualitatively similar even though the well controls obtained with the two models are quite different.

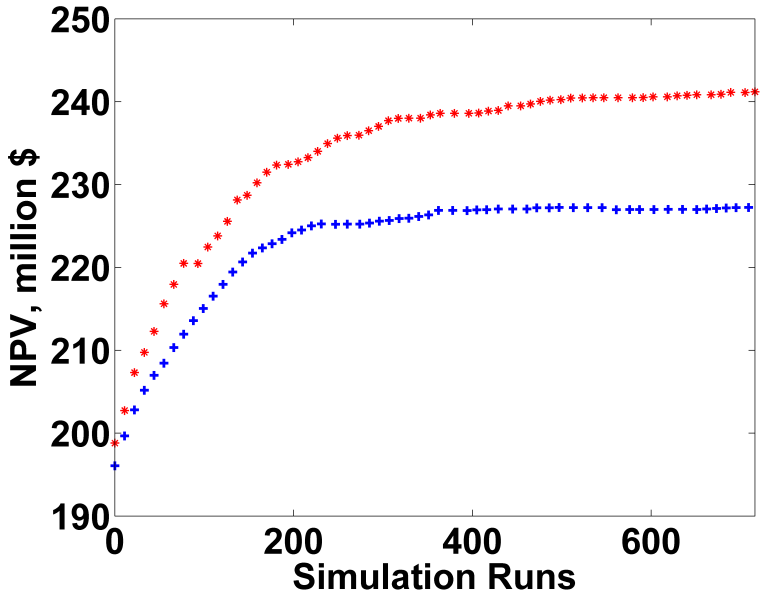
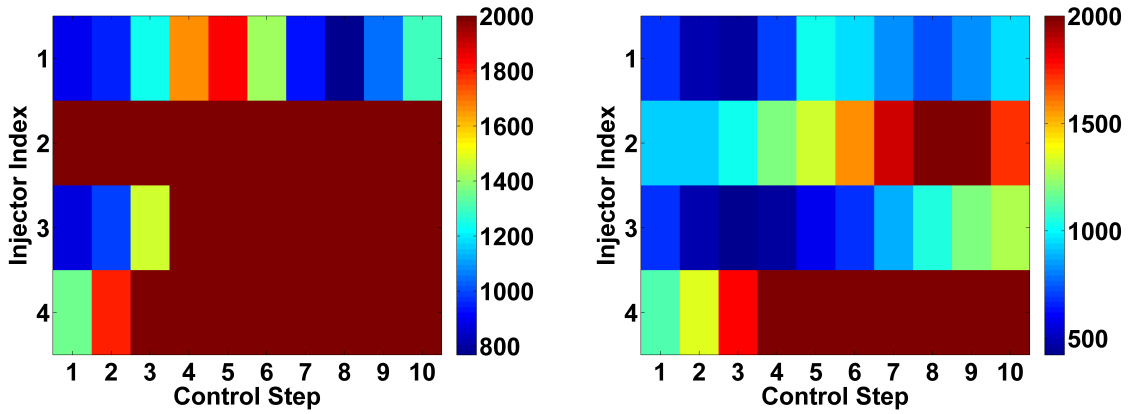


Figure 3.6: NPV versus simulation runs, channelized reservoir. Red stars denote the independent optimization with Eclipse; blue + denote the optimization with INSIM-FT.

3.5.3 Field Example with Aquifer

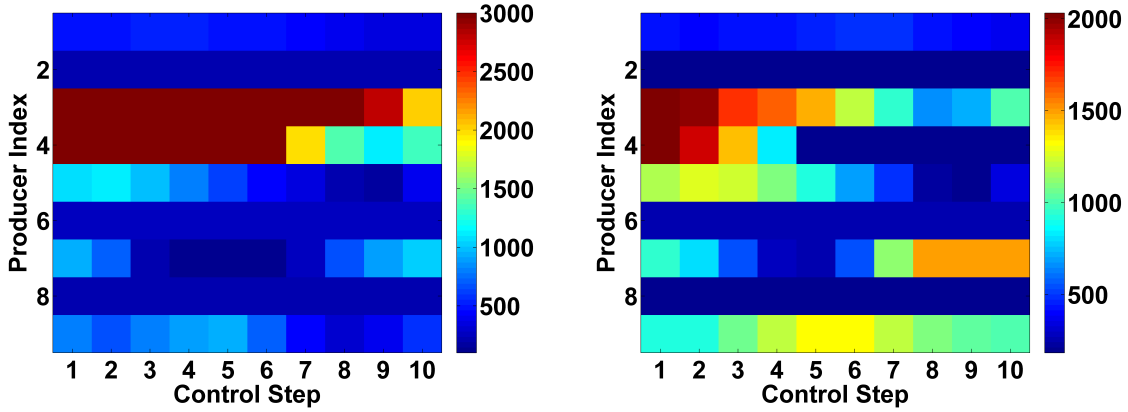
As shown in Chapter 2, the reservoir for this field example has 13 producing wells and no injection wells. The pressure is maintained by a constant-pressure aquifer. Based on the



(a) Optimal well controls for injectors with INSIM-FT model

(b) Optimal well controls for injectors with Eclipse simulation model

Figure 3.7: Estimated optimal water injection rate for injectors.



(a) Optimal well controls for producers with INSIM-FT model

(b) Optimal well controls for producers with Eclipse simulation model

Figure 3.8: Estimated optimal BHP for producers at different control steps.

Table 3.5: NPV comparison between INSIM-FT and Eclipse validations. INSIM-FT initial and optimal, respectively, represent the NPVs generated with INSIM-FT using the initial and optimal INSIM-FT well controls. Eclipse V1 inital and optimal, respectively, represent the results obtained by running Eclipse with the same initial and optimal well controls from INSIM-FT; Eclipse V2 initial denotes the NPV generated using the initial well controls used for INSIM-FT in the true Eclipse simulation model. Eclipse V2 optimal denotes the results by doing optimization directly on the reservoir simulation model (the truth).

	INSIM-FT		Eclipse V1		Eclipse V2	
	Initial	Optimal	Initial	Optimal	Initial	Optimal
NPV, million USD	197.07	227.99	198.80	234.04	198.80	241.28

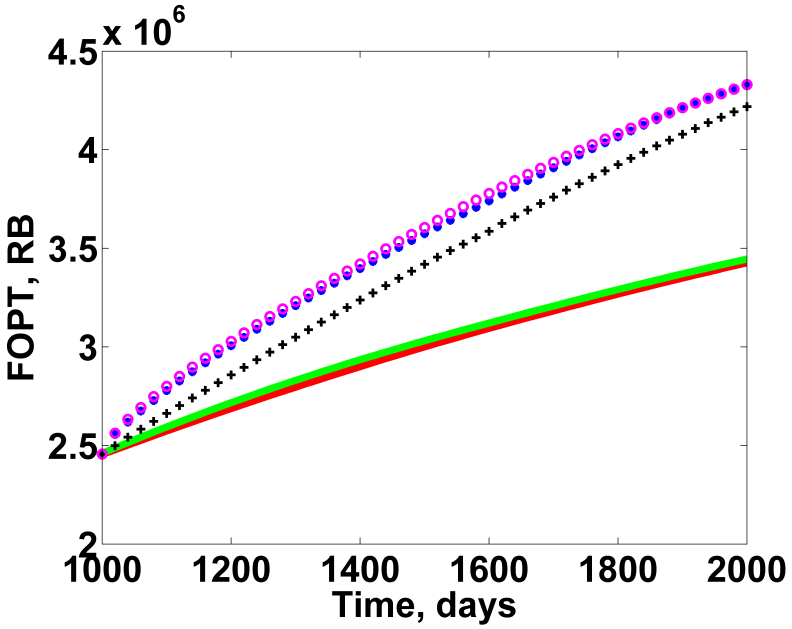
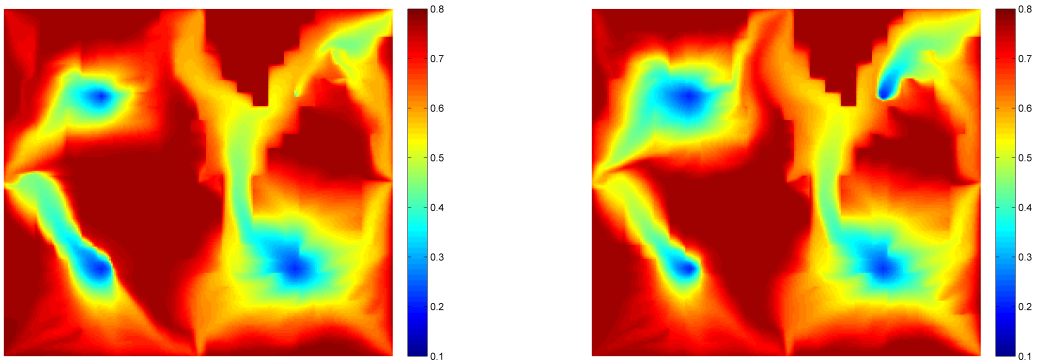


Figure 3.9: Cumulative oil production versus time, Example 2. Green curve represents cumulative oil production obtained from the initial guess of optimal well controls using INSIM-FT as the forward model; red curve represents the cumulative oil production from Eclipse with the initial guess of optimal well controls; black pluses denote the results calculated with the optimized INSIM-FT control using the INSIM-FT forward model; blue circles denote the results from Eclipse true model using the optimal controls estimated with INSIM-FT; pink circles denote the results from independent optimization by Eclipse.



(a) INSIM-FT (b) Eclipse

Figure 3.10: Oil saturation distributions obtained at the end of production life by applying the optimal well controls of INSIM-FT and Eclipse, channelized reservoir.

well locations, the connection map is built as shown in Fig. 2.24. Between each connected pair of producers, an imaginary well is placed at the middle point of the connection to form two connections and each imaginary well is also connected to the virtual node representing the

Table 3.6: Optimal NPVs obtained with INSIM-FT and Eclipse with two different initial conditions of well controls. INSIM-FT represents the optimal NPV generated with INSIM-FT; Eclipse V1 represent the NPV obtained by running Eclipse with the optimal well controls from INSIM-FT; Eclipse V2 denotes the NPV by optimizing directly on the reservoir simulation model (the truth), channelized reservoir.

	Injection Rate	Producer BHP	INSIM-FT	Eclipse V1	Eclipse V2
Case 1	600 RB/day	1,040 psi	235×10^6 USD	240×10^6 USD	247×10^6 USD
Case 2	1,400 RB/day	1,800 psi	220×10^6 USD	229×10^6 USD	235×10^6 USD

aquifer. The 2,050 days' producing period is history matched and production optimization is done for the next 2,010 days. Due to the lack of historical BHP information, we have no way to use BHP control for optimization. Therefore, the well controls for production optimization are the total production rate with initial controls being the controls at the end of the historical period. The total production rates have an upper bound of 1.2 times the maximum rate and a lower bound of 0.1 times the minimum rate obtained during production history. The optimization time interval contains nine control steps with the length of every control step equal to 238 days. The correlation length for temporal smoothness is 714 days. The control variables are total production rates of producing wells. The number of perturbations used for gradient estimation is 10; the maximum number of stepsize cuts is 5; the initial stepsize for each steepest ascent direction is 0.1; the maximum number of allowable simulation runs, N_{maxsim}, is 1,000; the maximum number of optimization iterations, $N_{maxiter}$, is 60. The steepest ascent algorithm terminates if both Eqs. 3.20 and 3.21 are satisfied or if N_{maxim} or $N_{maxiter}$ is reached. In this example, it takes less than 5 minutes to finish the production optimization with INSIM-FT as the forward model.

The initial well controls are shown in Fig. 3.11 and the optimal producing rates obtained from production optimization with INSIM-FT as the forward model are shown in Fig. 3.12. The NPV versus simulation runs is shown in Fig. 3.13. During the optimization run, the NPV increases from 557 million USD to 782 million USD in 700 INSIM-FT runs.

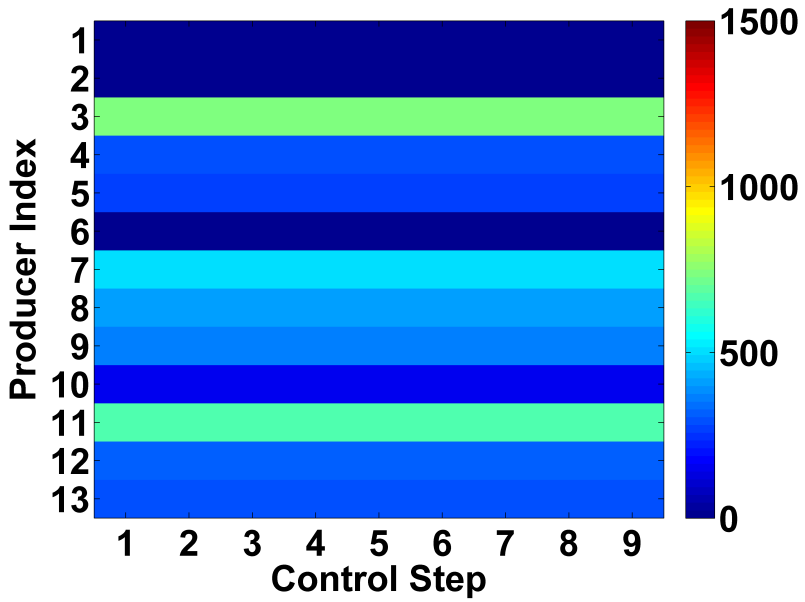


Figure 3.11: Initial guess of optimal total liquid production rate.

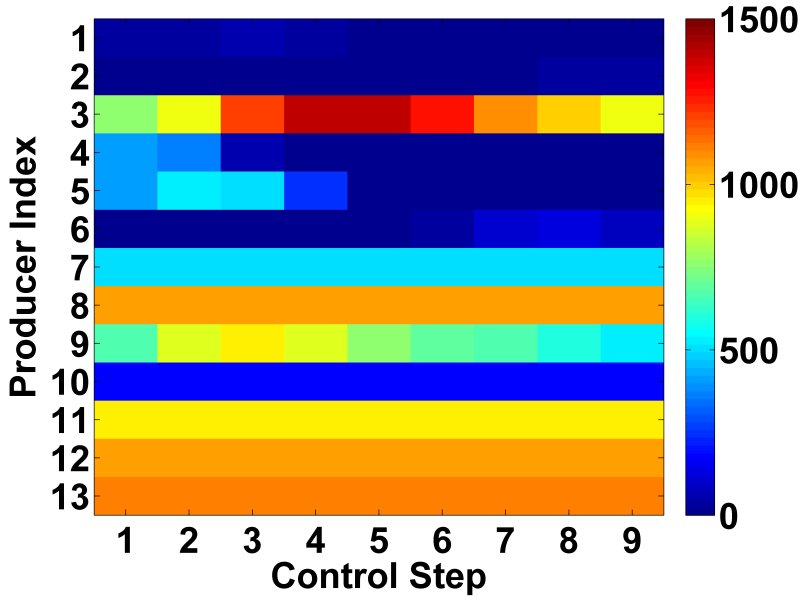


Figure 3.12: Optimal liquid production rate obtained with INSIM-FT.

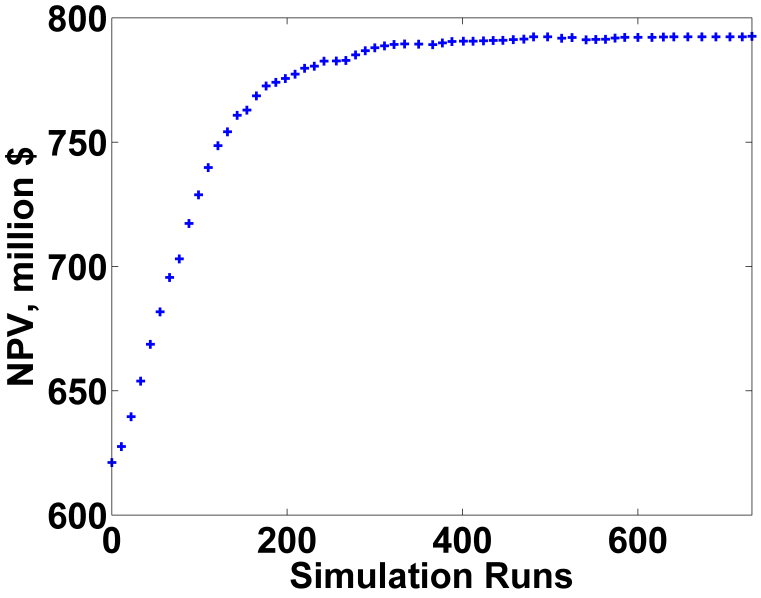


Figure 3.13: NPV versus INSIM-FT simulation runs.

CHAPTER 4

INSIM-FT IN THREE-DIMENSIONS WITH GRAVITY

In Chapter 2, we developed a simple data-driven model, INSIM-FT for reservoir flow and transport as a replacement for a reservoir simulator. However, INSIM-FT only considers two-dimensional flow in a reservoir only allows vertical wells. In this chapter, we develop a new data-driven model that extends the interwell numerical simulation model with front-tracking (INSIM-FT) from single-layer reservoirs to full three-dimensional (3D) multi-layer reservoirs. The new model, which is referred to as INSIM-FT-3D, can be used for history matching and reservoir performance predictions for a three-dimensional reservoir under waterflooding. The novelty of the new approach includes four points. First, INSIM-FT-3D replaces the original Riemann solver in INSIM-FT by a new Riemann solver based on a convex-hull method that enables the solution of the Buckley-Leverett problem with gravity, where a fractional flow function may have more than one inflection point. Second, unlike the original INSIM-FT model, which assumes all wells are vertical, the INSIM-FT-3D model allows for the inclusion of wells with arbitrary trajectories with multiple perforations. To do so, INSIM-FT-3D includes the well index of well performations (perforated zones) as history-matching parameters. Third, INSIM-FT-3D applies Mitchell's best-candidate algorithm to automatically generate the imaginary wells that are evenly distributed in the reservoir given a set of prefixed actual well nodes and fourth INSIM-FT-3D utilizes our own modification of Delaunay triangulation to build the 3D connection map necessary to use the general INSIM-FT-3D formulation.

History matching of the INSIM-FT-3D model parameters is performed using the ensemble-smoother with multiple data assimilations. The model parameters for history matching include the connection-based parameters, the parameters defining the power-law

relative permeabilities and the parameters defining the well indices for the perforated zones of a well with multiple perforated well segments.

The semi-analytical saturation solution from INSIM-FT-3D is validated by comparing the INSIM-FT-3D results with those obtained with an Eclipse reservoir simulation model on a fine grid where the observed data history-matched with INSIM-FT-3D were generated from Eclipse using the "true" reservoir simulation model. INSIM-FT-3D is applied for history matching and prediction for a multi-layer synthetic reservoir with a complex channel system, a real field example and a large-scale synthetic field example.

4.1 Methodology

INSIM-FT introduced in Chapter 2 is an two-dimensional (2D) numerical simulation model for a two-phase water oil system which characterizes a reservoir as a set of one dimensional (1D) connective flow volumes, which are similar to stream tubes. A schematic representation of a simple case is shown in Fig. 4.1, where each red circle represents the static bulk volume controlled by the well at its center and the dark gray areas represent the static constant bulk volumes of the connective flow units between pairs of well nodes. $V_{p,i,j}$ and $T_{i,j}$, respectively, represent the pore volume and transmissibility between connection (i, j). These set of $T_{i,j}$'s and the set of $V_{p,i,j}$'s are parameters for history matching with INSIM-FT. In addition, relative permeability curves are included as model parameters in the INSIM-FT model.

In this chapter, we develop a new data-driven model for 3D reservoirs with gravity based on the original INSIM-FT model proposed in Chapter 2. The 3D model still uses a set of 1D well interconnected volumes as shown in Fig. 4.1 but the line connecting a pair of nodes is not necessarily horizontal and thus, gravity affects the flow through the pore volume connecting the nodes. The model parameters for INSIM-FT are also used in INSIM-FT-3D. Like INSIM-FT, INSIM-FT-3D is a rate-control-based model. For history matching with INSIM-FT-3D, the total rate of injectors and producers must be specified at each time step and the oil rate is the observed data to be history matched using ES-MDA

(Emerick and Reynolds, 2012, 2013a,b). In the following context, the key formulations for the INSIM-FT-3D model are given below where field units are utilized throughout.

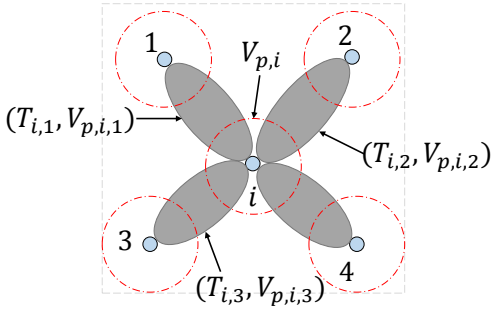


Figure 4.1: Connective units between wells of INSIM.

4.1.1 Pressure Equation

Following the derivation in Chapter 2, the total mobility for oil and water phase (λ_t) at a given water saturation, S_w, is defined by

$$\lambda_t(S_w) = \lambda_w(S_w) + \lambda_o(S_w), \tag{4.1}$$

where

$$\lambda_w(S_w) = \frac{k_{rw}(S_w)}{\mu_w} \tag{4.2}$$

is water mobility, and

$$\lambda_o(S_w) = \frac{k_{ro}(S_w)}{\mu_o} \tag{4.3}$$

is the oil mobility, μ_m, $m = o, w$ denote the oil and water viscosities in cp which are assumed to be constant and known and k_{rm}, $m = o, w$ represent the oil and water relative permeabilities, respectively. The general power-law relative permeability functions assumed here are given by Eqs. 2.17 through 2.19.

As for the two-dimensional version of INSIM-FT, we assume that, in Eqs. 2.17 through 2.19, S_{iw} and S_{or} for INSIM-FT-3D are known accurately but a, n_w and n_o are included as parameters to be estimated by history matching. The transmissibility for connection (i, j) is

defined in Eq. 2.2. The total compressibility at control volume i at time t^n, $c_{t,i}$ is defined in Eq. 2.3 as used in INSIM-FT and we assume throughout that the rock compressibility (c_r) and fluid compressibilities ($c_m,\ m = o, w$) in psi^{-1} are known constants.

For now, we assume that all the wells only have one perforated well segment in the reservoir. (Later, we will show how to handle wells with more than one perforated segment.) We consider gravity effects but assume capillarity effects are negligible. By combining the oil and water discretized flow equations, following the well-known IMPES formulation (Aziz and Settari, 1979), the pressure equation for control volume i in INSIM-FT considering gravity effects is given by

$$\sum_{j=1}^{n_{c,i}} T_{i,j}^{n-1}\Big(p_j^n - p_i^n - \big[(\gamma_{o,i,j}^{n-1}\frac{\lambda_o(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})} + \gamma_{w,i,j}^{n-1}\frac{\lambda_w(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})})(D_j - D_i)\big]\Big) + q_{t,i}^n = \frac{1}{5.615}\frac{c_{t,i}^{n-1}V_{p,i}^{n-1}}{\Delta t_n}(p_i^n - p_i^{n-1}), \quad (4.4)$$

where the superscript n denotes the nth time step; $n_{c,i}$ represents the number of volumes connected directly to volume i; D_i and D_j, respectively, represent the depth of the center of nodes i and j in ft; $q_{t,i}$ in RB/day is the total rate of the well in volume i, where $q_{t,i}$ is positive for injection and negative for production; $\gamma_{m,i,j}^{n-1},\ m = o, w$, denote the average specific weights of oil and water in psi/ft^3 for connection (i, j) at time level $n-1$ given by

$$\gamma_{m,i,j}^{n-1} = 0.5(\gamma_{m,i}^{n-1} + \gamma_{m,j}^{n-1}),\ m = o, w, \quad (4.5)$$

where

$$\gamma_{m,i}^{n-1} = \gamma_{m,i}^0(1 + c_m(p_i^{n-1} - p^0)), \quad (4.6)$$

and

$$\gamma_{m,j}^{n-1} = \gamma_{m,j}^0(1 + c_m(p_j^{n-1} - p^0)), \quad (4.7)$$

where p^0 is the initial reservoir pressure, which is assumed known at datum and varies only with depth. From Eq. 2.2, $T_{i,j}^{n-1}$ in Eq. 4.4, is given by

$$T_{i,j}^{n-1} = 1.127 \times 10^{-3} k_{i,j} A_{i,j} \frac{\lambda_t(S_{w,i,j}^{n-1})}{L_{ij}} = T_{i,j}^{0} \frac{\lambda_t(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{0})}, \tag{4.8}$$

where $\lambda_{t,i,j}^{0}$ is specified as the oil mobility at the initial water saturation, which is assumed to be equal to irreducible water saturation. Following Chapter 2, $V_{p,i}^{n-1}$ in Eq. 4.4 is required to satisfy

$$V_{p,i}^{n-1} = 0.5 \sum_{j=1}^{n_{c,i}} V_{p,i,j}^{n-1}. \tag{4.9}$$

Note that Eq. 4.9 ensures that if the $V_{p,i}$'s sum to the total reservoir pore volume, then $V_{p,i,j}$'s also sum to the total pore volume. $V_{p,i,j}^{n-1}$ is approximated by the first-order Taylor series expansion given in Eq. 2.8. Assuming all the connections use the same set of relative permeabilities, the preceding formulation indicates that the only unknown parameters in Eq. 4.4 are $T_{i,j}^{0}$'s, $V_{p,i,j}^{0}$'s, a, n_w and n_o.

4.1.2 Wells with Multiple Perforated Segments

Eq. 4.4, which is a material balance equation, assumes that $q_{t,i}^{n}$ is the average liquid flow rate at the nth time step produced from (or injected into) control volume $V_{p,i}$ via the section of a well that is perforated in volume $V_{p,i}$. However, this well could also be perforated in other control volumes and in this case, $q_{t,i}^{n}$ is not the total rate of production (or injection) of the well. If volume i represents an imaginary well, Eq. 4.4 with a zero rate is also valid for the control volume controlled by the imaginary well. Here, we extend Eq. 4.4 to the case where a well, w, has n_{wc}^{w} perforated well segments, each associated with a different control volume. By considering each perforated well segment as a node in INSIM-FT-3D, the control volume associated with the node representing the kth perforated segment of well w is given the index i_k and is denoted by V_{p,i_k}, for $k = 1, 2, \ldots, n_{wc}^{w}$, where n_{wc}^{w} is the number of perforated segments of well w, the pressure governing equation for the control volume V_{p,i_k} is given by

$$\sum_{j=1}^{n_{c,i_k}} T_{i_k,j}^{n-1}\Big(p_j^n - p_{i_k}^n - \big[(\gamma_{o,i_k,j}^{n-1}\frac{\lambda_o(S_{w,i_k,j}^{n-1})}{\lambda_t(S_{w,i_k,j}^{n-1})} + \gamma_{w,i_k,j}^{n-1}\frac{\lambda_w(S_{w,i_k,j}^{n-1})}{\lambda_t(S_{w,i_k,j}^{n-1})})(D_j - D_{i_k})\big]\Big)$$
$$+ q_{t,i_k}^n = \frac{1}{5.615}\frac{c_{t,i_k}^{n-1}V_{p,i_k}^{n-1}}{\Delta t_n}(p_{i_k}^n - p_{i_k}^{n-1}), \quad (4.10)$$

where n_{c,i_k} is the number of volumes connected to volume i_k; $T_{i_k,j}^{n-1}$ is the transmissibility between volume i_k and j, where j now can be a perforated volume; $p_{i_k}^n$ is the average pressure for control volume i_k at time level n; $S_{w,i_k,j}$ is the upstream water saturation between volume i_k and j; q_{t,i_k}^n is the total rate in RB/Day that flow in/out the kth control volume, V_{p,i_k}, of well w with q_{t,i_k}^n positive for injection and negative for production, and D_{i_k} is the depth of the center of the kth segment of well w. By using an inflow performance relationship, q_{t,i_k}^n is

$$q_{t,i_k}^n = \mathrm{WI}_{i_k}^w \lambda_t(S_{w,i_k}^{n-1})(p_{\mathrm{wf}}^n + H_{w,k}^{n-1} - p_{i_k}^n), \quad (4.11)$$

where $\mathrm{WI}_{i_k}^w$ is the well index between the wellbore associated with the kth segment of well w and the reservoir volume V_{p,i_k}, which is not known accurately and has to be history matched. S_{w,i_k} is the water saturation of volume i_k at time level $n-1$; p_{wf} is the bottom-hole pressure of well w; $H_{w,k}$ is the well-bore pressure difference between the center of the kth segment of well w and the bottom-hole datum depth. We assume that the fluid density inside a wellbore is uniform with depth, and oil and water densities in the wellbore, respectively, are equal to the oil and water densities at the standard conditions. Then,

$$H_{w,k}^{n-1} = \frac{\gamma_{o,\mathrm{sc}}q_o^{n-1} + \gamma_{w,\mathrm{sc}}q_w^{n-1}}{q_o^{n-1} + q_w^{n-1}}(D_{i_k} - D_{\mathrm{bh}}), \quad (4.12)$$

where the subscript sc refers to the standard conditions of fluids; D_{bh} is the bottom-hole datum depth; q_o and q_w, respectively, represent the total flow rate of oil and the total flow rate of water at well w. In order to compute p_{wf}^n and $p_{i_k}^n$, $k = 1, 2, \ldots, n_{wc}^w$ simultaneously, we use the fact that the total flow rate from the well must be equal to the sum of rates through

all the perforated segments, i.e.,

$$q_{t,w}^n = \sum_{k=1}^{n_{wc}^w} q_{t,i_k}^n, \tag{4.13}$$

where $q_{t,w}^n$ is the specified total rate of well w at time level n. Note that $\mathrm{WI}_{i_k}^w$, $k = 1, 2, \ldots, n_{\mathrm{wc}}$ in Eq. 4.11 is not known a priori; therefore, these terms are included as history-matching parameters when using INSIM-FT-3D as the forward model. The total oil rate for well w at time level n, $q_{o,w}^n$, is computed as

$$q_{o,w}^n = \sum_{k=1}^{n_{wc}^w} \Big[\mathrm{WI}_{i_k}^w \lambda_o(S_{w,i_k}^{n-1})(p_{\mathrm{wf}}^n + H_{w,k}^{n-1} - p_{i_k}^n)\Big], \tag{4.14}$$

and the water rate $q_{w,w}^n$, is given by

$$q_{w,w}^n = \sum_{k=1}^{n_{wc}^w} \Big[\mathrm{WI}_{i_k}^w \lambda_w(S_{w,i_k}^{n-1})(p_{\mathrm{wf}}^n + H_{w,k}^{n-1} - p_{i_k}^n)\Big]. \tag{4.15}$$

4.1.3 Saturation Equation

By solving the linear system consisting of the pressure equations given by Eq. 4.4 for wells penetrating a single layer (and imaginary wells), or Eqs. 4.10 and 4.13 for wells with completions in multiple control volumes, the pressure of each node at time t^n is obtained. The total liquid flow rate along the connection (i, j) is computed by

$$q_{t,i,j}^n = T_{i,j}^{n-1}\Big(p_j^n - p_i^n - (\gamma_{o,i,j}^{n-1}\frac{\lambda_o(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})} + \gamma_{w,i,j}^{n-1}\frac{\lambda_w(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})})(D_j - D_i)\Big), \tag{4.16}$$

where a positive value of $q_{t,i,j}$ means node j is upstream of node i and a negative value means node i is upstream of node j. For the calculation of water saturation, we assume incompressible flow, so the partial differential equation governing the saturation distribution for linear flow through one of the connective pore volumes between a pair of nodes is in the

form of Buckley-Leverett equation given by

$$\frac{\partial S_w(x,t)}{\partial t} + \frac{5.615 q_{t,i,j}(t)}{\phi_{i,j} A_{i,j}} \frac{\partial f_w(x,t)}{\partial x} = 0 \quad \text{for} \quad 0 \le x \le L_{i,j}, \quad t^{n-1} \le t \le t^n; \tag{4.17}$$

where $q_{t,i,j}(t)$ is assumed equal to $q^n_{t,i,j}$ from t^{n-1} to t^n; f_w is the water fractional flow which includes gravitational effects but the capillarity effect is neglected; $\phi_{i,j}$ is the average porosity of the connective volume between wells i and j. The x direction, which varies from connection volume to connection volume and is in the direction of the line segment between node i and j: if $q_{t,i,j} > 0$, the direction is from j to i; otherwise the direction is from i to j. In Eq. 4.17, $\phi_{i,j} A_{i,j}$ is not assumed known but can be expressed as

$$\phi_{i,j} A_{i,j} = \frac{V_{p,i,j}}{L_{i,j}}. \tag{4.18}$$

Substituting Eq. 4.18 into Eq. 4.17 and setting $q_{t,i,j}(t) = q^n_{t,i,j}$ for $t^{n-1} \le t \le t^n$ yields

$$\frac{\partial S_w(x,t)}{\partial t} + \frac{5.615 q^n_{t,i,j} L_{i,j}}{V^n_{p,i,j}} \frac{\partial f_w(x,t)}{\partial x} = 0 \quad \text{for} \quad 0 \le x \le L_{i,j}, \quad t^{n-1} \le t \le t^n, \tag{4.19}$$

where $L_{i,j}$ is the known length of the line segment connecting node i to node j and the $V^n_{p,i,j}$ are related to $V^0_{p,i,j}$, which are parameters to be estimated by data assimilation (history matching). The water fractional flow function, f_w, in Eq. 4.17 is given by

$$f_w(S_w) = \frac{1 - 1.127 \times 10^{-3} \frac{k_{i,j} \lambda_o(S_w) A_{i,j}}{q^n_{t,i,j}} (\gamma^{n-1}_{w,i,j} - \gamma^{n-1}_{o,i,j}) \frac{D_i - D_j}{L_{i,j}}}{1 + \frac{k_{ro}(S_w)}{\mu_o} \frac{\mu_w}{k_{rw}(S_w)}}. \tag{4.20}$$

Substituting Eq. 2.2 into Eq. 4.20 yields

$$f_w(S_w) = \frac{1 - \frac{T^0_{i,j} \lambda_o(S_w)}{q^n_{t,i,j} \lambda_t(S^0_{w,i,j})} (\gamma^{n-1}_{w,i,j} - \gamma^{n-1}_{o,i,j})(D_i - D_j)}{1 + \frac{k_{ro}(S_w)}{\mu_o} \frac{\mu_w}{k_{rw}(S_w)}}. \tag{4.21}$$

To run INSIM-FT-3D as a forward model for history matching, we need to specify the

values of $V^0_{p,i,j}$ and $T^0_{i,j}$ for each connection, the global relative permeability functions, and the well index (WI) for each perforation of the wells with more than one perforated segment. For each INSIM-FT-3D step, we first solve the pressure equations (Eq. 4.4 for wells with single perforation and Eqs. 4.10 and 4.11 for wells with multiple perforations) for the pressure of each node implicitly by considering all the saturation-related terms are calculated with the saturation values from the previous time step. Then, the total liquid rate along each connection is computed using Eq. 4.16, and the saturation equation of Eq. 4.19 is analytically solved to obtain the saturation profile of each connective unit and the saturation value at each node. The procedure is repeated until we reach the last INSIM-FT-3D step. The method to analytically solve the saturation equation will be discussed in the next subsection.

4.1.4 Convex-Hull Method

Eq. 4.19 is generally a Cauchy problem for any given initial condition of water saturation. In Chapter 2, we used a front-tracking method to solve the Cauchy problem by splitting the problem into a sequence of sub-Riemann problems, which can be solved using a Riemann solver (Juanes and Patzek, 2003). The Riemann solver used in Chapter 2 assumes that the fractional flow function in Eq. 4.21 has one and only has one inflection point, i.e. is "S" shaped. However, the fractional flow function given in Eq. 4.21 can have more than one inflection point if the magnitude of the gravitational part of the fractional flow function is relatively large compared to the convective part (see Fig. 4.2); for a non S-shaped fractional flow function, the Riemann solver utilized in Chapter 2 cannot be applied. Thus, in order to solve the Riemann problems associated with the Cauchy problem in Eq. 4.19, we apply a more generic solution in INSIM-FT-3D; this solution procedure, which is referred to as the convex-hull method, can solve the Riemann problem given any shape of a fractional flow function (Holden et al., 1988). The convex hull of a set of points, X, in a plane or in a space is the smallest convex set that contains X. When X is a bounded subset of the plane, the convex hull may be visualized as the shape enclosed by a rubber band stretched around X (De Berg et al., 2000). Fig. 4.3 shows an example of the convex hull for a set X of points

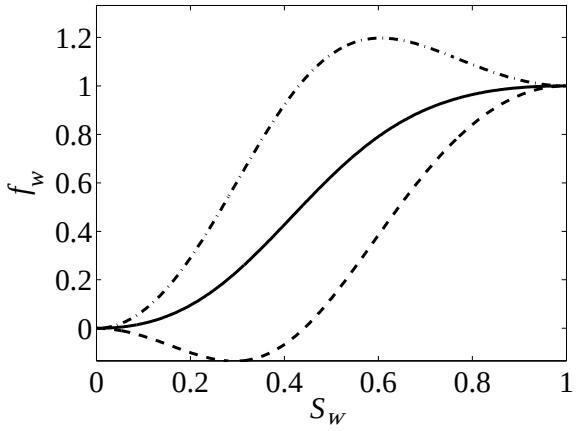


Figure 4.2: The fractional flow function with gravity included. The top dashed curve represents the fractional flow function for a vertical downwards flow; the solid curve represents the fractional flow function for horizontal flow when gravity has no effect; the bottom dashed curve represents the fractional flow function for vertical upwards flow.

in a 2D plane, where the upper part of the convex hull (upper hull) is the concave piecewise function in red and the lower part of the convex hull (lower hull) is the convex piecewise function in blue. The details for the computation of the convex-hull solution are described below.

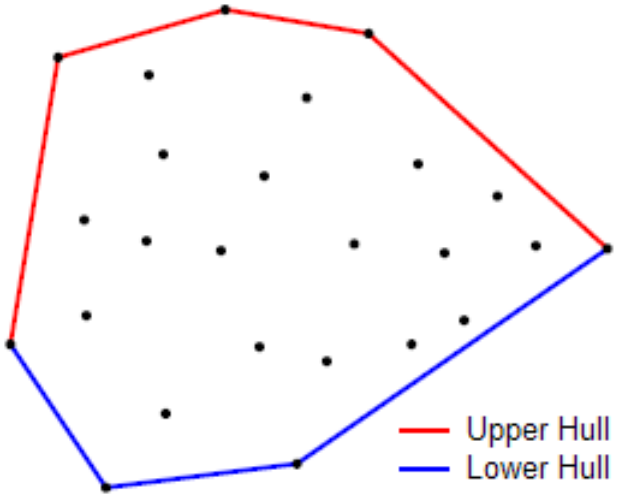


Figure 4.3: The convex hull of the set of points shown.

We consider a canonical Riemann problem (Juanes and Patzek, 2003) that consists

of finding a (weak) solution to the initial value problem:

$$\frac{\partial S_w(x,t)}{\partial t} + \eta \frac{\partial f_w(x,t)}{\partial x} = 0,$$
$$S_w(x,0) = \begin{cases} S_{wl} & \text{if } x < x_0 \\ S_{wr} & \text{if } x > x_0 \end{cases}, \tag{4.22}$$

where η is a constant and the partial differential equation (PDE) represents a conservation law. Assuming that $S_{wl} \neq S_{wr}$, the initial condition (saturation function at time zero) has a single discontinuity at x_0. The solution to the problem is a wave connecting the left (S_{wl}) and right (S_{wr}) states. To use the convex-hull method to solve the Riemann problem for any shape of the fractional flow function, the fractional flow function $f_w(S_w)$ in Eq. 4.22 must be approximated by a piecewise linear function, $f_{w,\text{pl}}(S_w)$ as shown in Fig. 4.4. Related to Fig. 4.4, $S_{w,\text{pl},1} < S_{w,\text{pl},2} < \ldots, < S_{w,\text{pl},N_\text{pl}}$ are the S_w coordinates of the endpoints that constitute the piecewise linear approximation of f_w on $[S_{w1}, S_{w2}]$, where $S_{w,\text{pl},1} = S_{w1}$ and $S_{w,\text{pl},N_\text{pl}} = S_{w2}$, and $f_{w,\text{pl}}(S_{w,\text{pl},j}) = f_w(S_{w,\text{pl},j})$, $j = 1, 2, \ldots, N_\text{pl}$. Though a Riemann problem can include any two states in the initial condition, for the purpose of illustration, we assume that S_{w1} and S_{w2} are the S_w coordinates of two points used to construct the piecewise linear approximation of the fractional flow curve in $f_{w,\text{pl}}(S_w)$ as shown in Fig. 4.4 and represent two states of a general Riemann problem, where $S_{w1} = S_{wl}$ and $S_{w2} = S_{wr}$ if $S_{wl} < S_{wr}$, and $S_{w1} = S_{wr}$ and $S_{w2} = S_{wl}$ if $S_{wr} < S_{wl}$. (If S_{w1} and S_{w2} are not associated with any endpoint in $f_{w,\text{pl}}(S_w)$, we simply add the two points $(S_{w1}, f_w(S_{w1}))$ and $(S_{w2}, f_w(S_{w2}))$ into the point set that comprises the piecewise linear fractional flow function, and then update $f_{w,\text{pl}}(S_w)$ before starting to solve Eq. 4.22.) The red curve above the fractional flow in Fig. 4.4 represents the upper part of the convex hull relative to the water saturation interval between S_{w1} and S_{w2}, and the blue curve represents the lower part of the convex hull for the saturation interval. The method we use to compute the convex hull is given in Appendix D.

The case where $S_{wl} > S_{wr}$ is considered first. Let $f_{w,u}(S_w)$ be the function corresponding to the upper part of the convex hull of $f_{w,\text{pl}}(S_w)$ relative to the saturation interval

between S_{wl} and S_{wr} (see Fig. 4.4), where $S_{wl} = S_{w2}$ and $S_{wr} = S_{w1}$. We denote the number of vertices (points) in the upper convex hull by N ($N \geq 2$). Note that N will always be equal to or smaller than number of endpoints, N_{pl}, of the piecewise linear approximation, $f_{w,\text{pl}}(S_w)$, for $S_{wr} \leq S_w \leq S_{wl}$, according to the definition of convex hull. We sort the S_w coordinates of vertices in the upper convex hull in an increasing order as $S_{w,j}$, $j = 1, 2, \ldots, N$,

$$S_{wr} = S_{w,1} < S_{w,2} < \ldots, S_{w,N} = S_{wl}, \tag{4.23}$$

and denote the point set corresponding to the upper convex hull by $\{(S_{w,j}, f_{w,u}(S_{w,j})), j = 1, 2, \ldots, N\}$. It is easy to see that $f_{w,u}(S_{w,j}) = f_{w,\text{pl}}(S_{w,j})$ for $j = 1, 2, \ldots, N$. Then, according to Holden et al. (1988), the exact solution of Eq. 4.22 is given by

$$S_w(x,t) = \begin{cases} S_{w,1}, & \text{if} \quad \dfrac{x - x_0}{t} \geq \eta \dfrac{f_{w,u}(S_{w,2}) - f_{w,u}(S_{w,1})}{S_{w,2} - S_{w,1}}, \\ \ldots, & \\ S_{w,j}, & \text{if} \quad \eta \dfrac{f_{w,u}(S_{w,j}) - f_{w,u}(S_{w,j-1})}{S_{w,j} - S_{w,j-1}} \geq \dfrac{x - x_0}{t} \geq \eta \dfrac{f_{w,u}(S_{w,j}) - f_{w,u}(S_{w,j+1})}{S_{w,j} - S_{w,j+1}}, \\ \ldots, & \\ S_{w,N}, & \text{if} \quad \dfrac{x - x_0}{t} \leq \eta \dfrac{f_{w,u}(S_{w,N}) - f_{w,u}(S_{w,N-1})}{S_{w,N} - S_{w,N-1}}, \end{cases} \tag{4.24}$$

where $j = 2, 3, \ldots, N-1$. Eq. 4.24 indicates the solution of Eq. 4.22 for $S_{wl} > S_{wr}$ is a set of shocks between S_{wl} and S_{wr}, and every two neighboring $S_{w,j}$'s, $j = 1, 2, \ldots, N$ in Eq. 4.23 are the saturations that define a shock. The equality of Eq. 4.24 holds because the shock wave (weak solution) contains an interval of water saturations representing a discontinuity

in the saturation profile. For the shock between the saturation interval of $(S_{w,j}, S_{w,j+1})$, $j = 1, 2, \ldots, N-1$, the shock speed, σ, is given by the Rankine-Hugoniot jump condition:

$$\sigma = \eta \frac{f_{w,u}(S_{w,j+1}) - f_{w,u}(S_{w,j})}{S_{w,j+1} - S_{w,j}}. \tag{4.25}$$

For the case where $S_{wl} < S_{wr}$, we denote the lower part of the convex hull by $f_{w,l}$ (see Fig. 4.4). Letting N be the number of vertices of this lower convex hull, we sort the x coordinates of vertices of $f_{w,l}(S_w)$ in an increasing order as $S_{w,j}$, $j = 1, 2, \ldots, N$, by

$$S_{wl} = S_{w,1} < S_{w,2} < \ldots, S_{w,N} = S_{wr}, \tag{4.26}$$

then as shown in Holden et al. (1988), the exact solution of Eq. 4.22 is given by

$$S_w(x,t) = \begin{cases} S_{w,1}, & \text{if} \quad \dfrac{x - x_0}{t} \leq \eta \dfrac{f_{w,l}(S_{w,2}) - f_{w,l}(S_{w,1})}{S_{w,2} - S_{w,1}}, \\ \ldots, & \\ S_{w,j}, & \text{if} \quad \eta \dfrac{f_{w,l}(S_{w,j}) - f_{w,l}(S_{w,j-1})}{S_{w,j} - S_{w,j-1}} \leq \dfrac{x - x_0}{t} \leq \eta \dfrac{f_{w,l}(S_{w,j}) - f_{w,l}(S_{w,j+1})}{S_{w,j} - S_{w,j+1}}, \\ \ldots, & \\ S_{w,N}, & \text{if} \quad \dfrac{x - x_0}{t} \geq \eta \dfrac{f_{w,l}(S_{w,N}) - f_{w,l}(S_{w,N-1})}{S_{w,N} - S_{w,N-1}}, \end{cases} \tag{4.27}$$

where $j = 2, 3, \ldots, N-1$. Eq. 4.27 indicates the solution of Eq. 4.22 for $S_{wr} > S_{wl}$ is a set of shocks between S_{wl} and S_{wr}, and every two neighboring $S_{w,j}$'s, $j = 1, 2, \ldots, N$ in Eq. 4.26 is a shock. For the shock between $(S_{w,j}, S_{w,j+1})$, $j = 1, 2, \ldots, N-1$, the shock speed, σ, is

given by the Rankine-Hugoniot jump condition:

$$\sigma = \eta \frac{f_{w,l}(S_{w,j+1}) - f_{w,l}(S_{w,j})}{S_{w,j+1} - S_{w,j}}. \tag{4.28}$$

Note that the above solution is useful only if we obtain the convex hull of f_w relative to any given interval between S_{wl} and S_{wr}. In our applications, Graham's scan (De Berg et al., 2000) is applied to compute the convex hull for a piecewise linear fractional flow function as shown in Appendix D. By applying the generic solution for the scalar form of the Riemann problem as described in Eqs. 4.24 and 4.27, Eq. 4.17 with any initial condition of water saturation can be solved with the same front-tracking method used in Chapter 2 for 2D INSIM-FT.

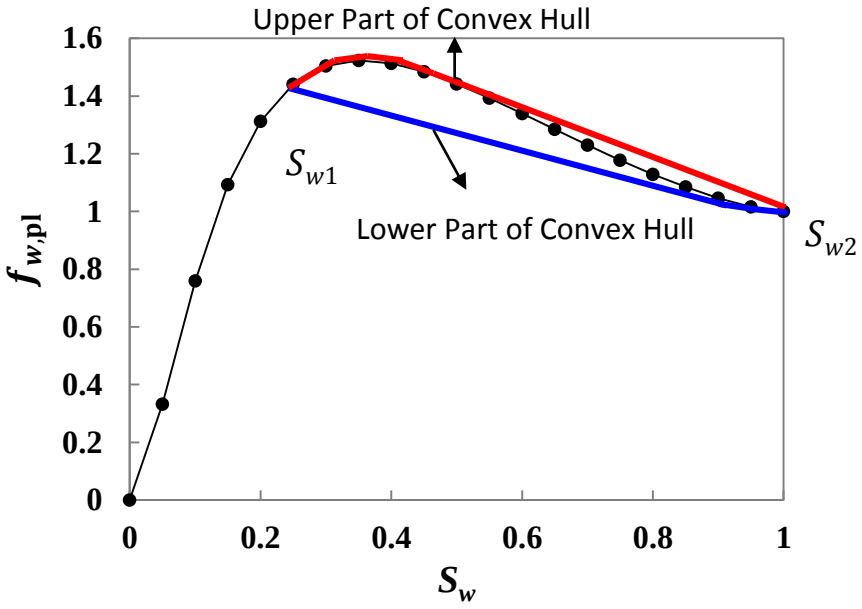


Figure 4.4: A piecewise linear approximation for a fractional flow function. The red curve and blue curve comprise the convex hull of the fractional flow function relative to a given saturation interval between S_{w1} and S_{w2}. The red curve is the upper part of the convex hull and the blue curve is the lower part of the convex hull.

4.1.5 Generate Imaginary Wells Using Mitchell's Best-Candidate Algorithm

Imaginary wells are introduced in the INSIM-FT model to add more flow paths between injectors and producers. Although some general instructions on how to place the

imaginary wells in a reservoir are provided in Chapter 2, e.g., adding imaginary wells between two adjacent injectors or two producers to prevent the direct connection between two injectors or producers, our heuristic method of inserting imaginary wells is not easy to apply to 3D problems which may include wells with arbitrary trajectories. Here, to add imaginary wells, we apply Mitchell's best candidate algorithm to sequentially place a number of imaginary well nodes inside a bounded reservoir given a number of fixed actual well nodes, so that each newly added imaginary well node is farthest away from all existing well nodes, i.e., farthest away from all the previously placed imaginary wells and fixed actual well nodes. This procedure tends to give a fairly even distribution of the imaginary wells. Mitchell's best-candidate algorithm (Mitchell, 1991) is described in Appendix E. One still has to determine a priori the number of imaginary well nodes, n_I, to be added and this may require some engineering judgment based on an inspection of the nodes generated. As a rule of thumb, we set n_I to be equal to one to two times the number of actual well nodes.

To illustrate the performance of Mitchell's best-candidate algorithm, a simple 2D example is presented next. Later, a 3D example is considered. As shown in Fig. 4.5, the square 2D reservoir has four horizontal wells and each well has three nodes. The length and width of this reservoir are one in dimensionless units. We apply Mitchell's best-candidate algorithm to add 16 imaginary wells to this reservoir and the results are shown in Fig. 4.6. Note that all the nodes, including the actual well nodes, are fairly uniformly distributed in the reservoir. Next, we consider the problem of generating the connection map.

4.1.6 Create a Connection Map Using Delaunay Triangulation with A Modification

Given a set of well nodes, a connection map that connects these nodes with a set of line segments are required for history matching with INSIM-FT (or INSIM-FT-3D). The set of line segments function effectively define a set of connecting pairs of wells for streamtube simulation models and water saturations are solved along each of the 1D connections; see the discussion of Fig. 2.1.

The original INSIM-FT model utilized a method (Zhao et al., 2015) that connects

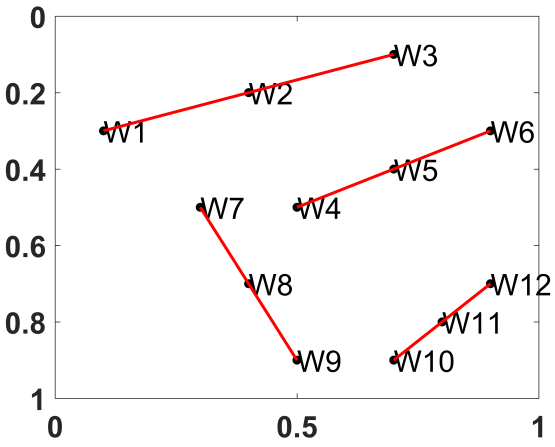


Figure 4.5: A 2D reservoir with four horizontal wells. The black dots represent the different perforated segments of wells and the red line segments represent the well trajectories.

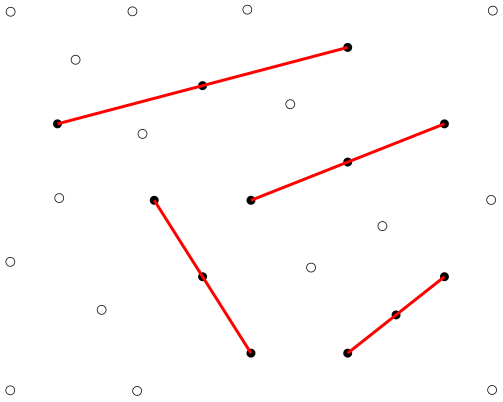


Figure 4.6: The imaginary well nodes generated with Mitchell's best candidate. The open circles represent the imaginary well nodes; red lines represent wells and the solid circles represent actual well nodes.

wells based on the distances between wells for a single layer reservoir and is not feasible for generating a connection map for a 3D model. Instead, to create a connection map for INSIM-FT-3D, Delaunay triangulation (Edelsbrunner et al., 1992) is applied with a small modification.

In computational geometry, Delaunay triangulation provides a method to triangulate a set of points in a d-dimensional space ($d \geq 2$). For a 2D problem, Delaunay triangulation of a given set X of points must satisfy the condition that no point in X is in the interior of the circumcircle of any triangle; this condition is called the "Delaunay condition." (The

circumcircle of a triangle is the circle that passes through the three vertices of the triangle.) There always exists a Delaunay triangulation for any set X of points in two dimensions and the triangulation is unique as long as no four or more points in X are cocircular (Devadoss and O'Rourke, 2011), i.e., we cannot find a circle so that four or more points in X lie over the perimeter of the circle. Delaunay triangulation maximizes the minimum angle of all the angles of the triangles in the triangulation (Devadoss and O'Rourke, 2011). Compared to any other triangulation of the points, the smallest angle in Delaunay triangulation is at least as large as the smallest angle in any other triangulation (Edelsbrunner et al., 1992). One well-known application of Delaunay triangulation is to generate a high-quality mesh for the finite element method. By high-quality, we mean that no triangles with extremely small (or large) angles are included in the triangulation.

The idea of using Delaunay triangulation to generate a connection map where the connection are the sides of all triangles is that given a set of well nodes in 2D or 3D space, Delaunay triangulation can always connect these nodes using triangles satisfying the "Delaunay condition," i.e., the requirement that the circumcircles of all triangles have empty interiors, i.e., no node lies in the interior of any circumcircle. Compared to other triangulation methods, Delaunay triangulation is more likely to generate the connection map that precludes connections of two nodes that are far apart, which is one of the rules for connecting wells in INSIM-FT. The quality of a Delaunay triangulation does depend on the distribution of the points. If the well nodes tend to be evenly placed, Delaunay triangulation tends to generate a high-quality connection map where all connections has a roughly the same length. Thus, placing the imaginary wells (nodes) so that are as uniformly placed as possible using the aforementioned Mitchell's best-candidate algorithm helps to create a better connection map when using Delaunay triangulation.

Delaunay triangulation is a convenient and automatic tool that enables the connection of nodes of wells when the wells are in an irregular pattern in three dimensions with well centerlines ranging from vertical to horizontal, and each well trajectory is represented by several well nodes along the well centerline. Delaunay triangulation eliminates the dif-

ficult work of manually constructing the connection map. For our applications, Delaunay triangulation implementation is from Matlab (MathWorks, 2004) software.

Delaunay triangulation for the node set of Fig. 4.6 is shown in Fig. 4.7. In Fig. 4.7, note that in regions close to the boundary of the reservoir, Delaunay triangulation still generates some unsatisfactory connections and corresponding triangles with large angles, e.g., for node pairs between (IM1,IM5), (IM1,IM12), (IM1,W9) and (W9,IM5). The reason is that at these boundary regions, there are not enough nodes to generate a "good" shape of the Delaunay triangles. Therefore, we should remove the connections with large angles in the connection map generated by Delaunay triangulation. For example, the connection between IM1 and IM5 in Fig. 4.7 will be eliminated by checking the interior angle between the edge (IM1,W9) and the edge (W9,IM5) of the triangle composed of the node IM1, W9 and IM5; if the angle is very large, e.g. greater than or equal to 120 degree, then IM1, W9 and IM5 are almost collinear and the connection between IM1 and IM5 should be removed, because IM1 and IM5 is not likely to be directly connected given the existence of W9. Following this idea, we check each interior angle of each Delaunay triangle and remove the connection (edge) corresponding to a interior angle that is greater than 120 degree. By applying this modification after triangulating the well nodes, the new connection map is shown in Fig. 4.8, where we can see that all the unsatisfactory connections with large angles are removed; compare the results of Fig. 4.8 with those of Fig. 4.7.

As noted in Chapter 2, when the INSIM-FT methodology is applied, a direct connection between a pair of injectors or producers may produce non-physical results. Here, we regulate the connection map generated with Delaunay triangulation by removing direct connections between two nodes (well segments), when both nodes are associated with injection wells or both nodes are associated with production wells. It should be strongly emphasized that nodes from the same well, cannot be connected to each other.

4.2 Example 1: Toy Problem

We consider a simple incompressible 1D reservoir to validate the analytical water

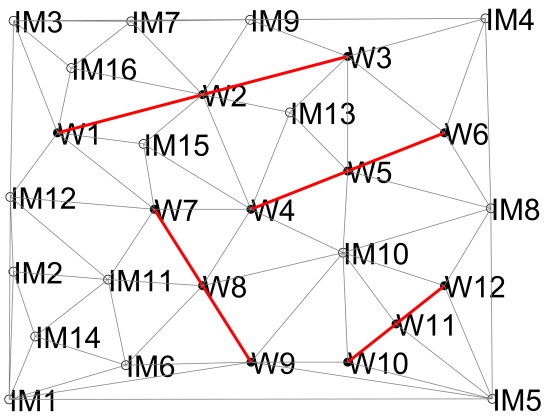


Figure 4.7: Delaunay triangulation for the node set of Fig. 4.6.

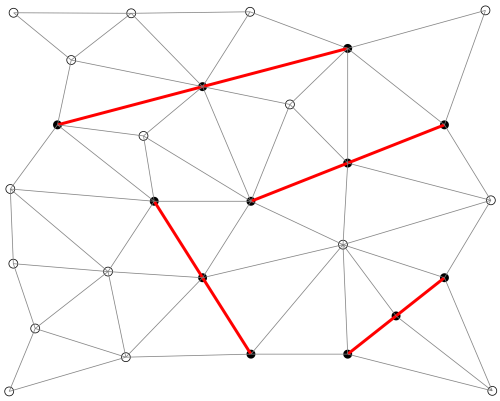


Figure 4.8: The connection map by applying a modification after Delaunay triangulation.

saturation results obtained from INSIM-FT-3D when gravity is included. We compare results with and without considering gravitational effects and set two different configurations for this 1D reservoir. The reservoir has dimensions of 100 ft by 100 ft by 1,000 ft, where the flow only happens along the z axis in order to include strong gravity effects. The corresponding Eclipse model has 1,000 gridblocks in the z direction and each gridblock has thickness of 1 ft. One water injector and one producer are placed in this vertical reservoir, where the injector is placed in the center of the top gridblock and the producer is placed in the center of the bottom gridblock to consider downward flow with gravity. To consider upward flow against the gravitational force, we simply switch locations of the two wells so the water injector is at the bottom and the producer is at the top. Then, we consider flow with gravity

excluded, i.e., rotate the reservoir so that flow is in the x direction and the gravitational force does not affect flow. The corresponding Eclipse model has 1,000 gridblocks of one-foot width in the x direction; the water injector is in the center of the left most gridblock and the producer is in the center of the right most gridblock. The water density is set to 62.428 lbm/ft^3 and oil density 56.94 lbm/ft^3. The porosity is 0.2. The compressibility for oil, water and rock, are set equal to 10^{-8} psi^{-1} to virtually eliminate the compressibility effects. The total water injection rate of the injector and the total production rate of the producer are set equal to 200 RB/day for all scenarios. The water-cut results obtained from INSIM-FT-3D and Eclipse for the producer under different scenarios, where flow can be horizontal, vertical downward or vertical upward, are shown in Fig. 4.9. As seen from Fig. 4.9, the INSIM-FT-3D water cut results are in very good agreement with the corresponding Eclipse results, which demonstrates that INSIM-FT-3D can correctly model gravitational effects.

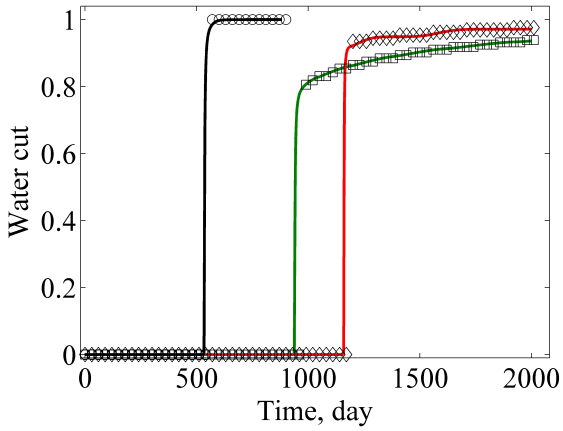


Figure 4.9: Water cut results from INSIM-FT compared with those from Eclipse; black solid curve represents the downward-flow results from Eclipse; dark green solid curve represents the horizontal flow results from Eclipse; red solid curve represents the upward flow results from Eclipse; INSIM-FT-3D results are shown open circles of same color as corresponding Eclipse results.

4.3 Example 2: Multilayer Channelized Reservoir

This example pertains to a six-layer reservoir which has three-facies, corresponding to a 3D channelized system, which is shown in Fig. 4.10. The true model is based on a six-layer CMG (IMEX, 2010) model with each layer having a 50 by 50 grid and the size of

each gridblock is 200 ft by 200 ft by 10 ft, where 10 ft is the thickness of each gridblock. Four injectors (I1-I4) and five producers (P1-P5) are placed in this reservoir with the locations shown in Fig. 4.11. The well trajectories are shown in Fig. 4.12, where the black solid lines represent the well trajectories of different wells. Note that I1, I2, P3 and P5 are vertical wells; I4 and P1 are inclined wells, and I3 and P4 are horizontal wells. Except for the horizontal wells, all the wells penetrate more than one layer of the reservoir. In Fig. 4.12, the red nodes on any particular black line segment (well trajectory) represent the "perforated" well segments of the associated well and the open circles represent 27 imaginary wells generated using the Mitchell's best-candidate algorithm. Fig. 4.13 shows the connection map generated using Delaunay triangulation modified by eliminating long connections and removing direct connections among any pair of injector nodes and any pair of producer nodes. The initial reservoir pressure is 3,800 psi. Other properties are defined in Table 4.1. As discussed in more detail later, synthetic production data for history matching with the INSIM-FT-3D model is obtained by running CMG with the true reservoir model where all injectors operate under a constant water injection rate of 10,000 RB/day and all producers operate at a constant BHP of 2,000 psi.

History matching is performed using ES-MDA (Emerick and Reynolds, 2012, 2013a,b) based on 200 initial sets of INSIM-FT-3D parameters, which in this example include the $T^0_{i,j}$'s, $V^0_{p,i,j}$'s for all connections, the well indices between wellbores and the reservoir at each perforated well segment, and the parameters defining a global set of relative permeabilities that is shared by all the interwell connections. The true values of parameters defining the relative permeability functions are given by

$$a_{\text{true}} = 0.6,\ n_{o,\text{true}} = 2,\ n_{w,\text{true}} = 2,\ S_{iw,\text{true}} = 0.1,\ S_{or,\text{true}} = 0.2.$$

To history match the relative permeabilities, S_{iw} and S_{or} are assumed to be known and a, n_w, n_o are estimated with the history-matching procedure. The mean values of a, n_o and

n_w are not equal to their true values, and more specifically, are given by

$$\overline{a} = 0.55,\ \overline{n}_o = 1.9,\ \overline{n}_w = 2.1.$$

The prior probability density functions (PDF's) of the three relative permeability parameters are specified by

$$\begin{aligned} &\mathcal{N}(\overline{a}, 0.05^2), \\ &\mathcal{N}(\overline{n}_o, 0.1^2), \end{aligned} \tag{4.29}$$

and

$$\mathcal{N}(\overline{n}_w, 0.1^2). \tag{4.30}$$

The prior ensemble for a, n_w, n_o is generated from sampling the normal distributions defined by Eqs. 4.29 and 4.30. For each connected well pair (i, j), to generate the initial ensemble of $V^0_{p,i,j}$, the mean value of $V^0_{p,i,j}$ is estimated as

$$\overline{V}^0_{p,i,j} = \frac{L_{i,j}}{\sum_{k=1}^{N_w-1}\sum_{l=k+1}^{N_w} L_{k,l}} \cdot V^0_{p,\text{tot}}, \tag{4.31}$$

where $\overline{V}^0_{p,i,j}$ represents the mean value of $V_{p,i,j}$. Then each prior ensemble member, $V^{0,k}_{p,i,j}$, for $k = 1, 2, \ldots, N_e$ is randomly generated by sampling the Gaussian distribution $\mathcal{N}(\overline{V}^0_{p,i,j}, (0.2\overline{V}^0_{p,i,j})^2)$. According to Eq. 2.2, for each connected well pair (i, j),

$$T^0_{i,j} = 1.127 \times 10^{-3} \frac{k_{i,j} V^0_{p,i,j} \lambda_t(S^0_{w,i,j})}{\phi^0_{i,j} L^2_{i,j}}. \tag{4.32}$$

Replacing $\phi^0_{i,j}$ in Eq. 4.32 by ϕ^0, where ϕ^0 is the estimated average porosity, and applying $\lambda_t(S^0_{w,i,j}) = \frac{1}{\mu_o}$ yields

$$T^0_{i,j} = 1.127 \times 10^{-3} \frac{k_{i,j} V^0_{p,i,j}}{\mu_o \phi^0 L^2_{i,j}}. \tag{4.33}$$

Letting $\overline{T}^0_{i,j}$ be the mean value of $T^0_{i,j}$, $\overline{T}^0_{i,j}$ is estimated by replacing $k_{i,j}$ in Eq. 4.33 by $\bar{k}$ (the

initial guess of the mean value of $k_{i,j}$) and is given by Eq. 4.34.

$$\overline{T}^0_{i,j} = 1.127 \times 10^{-3} \frac{\bar{k}\overline{V}^0_{p,i,j}}{\mu_o \phi^0 L^2_{i,j}}, \tag{4.34}$$

where $\bar{k} = 500$ md; $\phi^0 = 0.2$ and $\mu_o = 2.2$ cp. It is important to note that the mean values of the permeabilities of the three individual facies are 20 md, 200 md and 2,000 md. For $k = 1, 2, \ldots, N_e$, the prior ensemble member, $T^{0,k}_{i,j}$, is generated by sampling $\mathcal{N}(\overline{T}^0_{i,j}, (0.2\overline{T}^0_{i,j})^2)$. In order to generate the initial ensemble of well indices, the mean values of well indices are calculated from Peaceman's equation (Peaceman, 1978). For kth perforation of well w, the mean value of the well index is estimated by

$$\overline{\mathrm{WI}}^w_{i_k} = \frac{1}{141.2} \frac{\bar{k}h}{\ln(r_o/r_w) + s}, \tag{4.35}$$

where h is the thickness of the layer perforated; s is the skin factor and r_o is radius estimated by

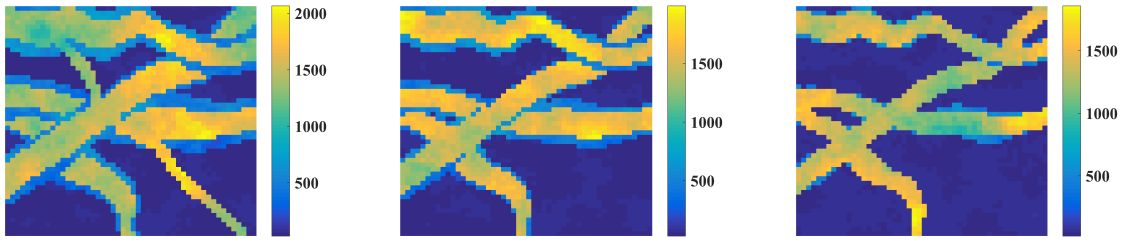
$$r_o = \sqrt{\frac{V^0_{p,i_k}}{\pi \phi^0 h}}, \tag{4.36}$$

where V^0_{p,i_k} is the pore volume of control volume i_k and is computed by Eq. 4.9. The initial ensemble of $\mathrm{WI}^w_{i_k}$ is generated by sampling $\mathcal{N}(\overline{\mathrm{WI}}^w_{i_k}, (0.2\overline{\mathrm{WI}}^w_{i_k})^2)$.

The historical period for history matching is 2,100 days and the subsequent 900 days is used for future performance predictions. The true total liquid rates for injectors and producers over the 3,000 days' production history are obtained by running the CMG true model and are specified in the INSIM-FT-3D model as the well operating conditions. The oil production rate, which is the observed data that is history matched, are generated by adding uncorrelated Gaussian random noise to the true data obtained from the CMG run with the true simulation model, where the standard deviation of the Gaussian noise is 10 RB/day.

The oil production rates for four representative wells from 200 prior realizations of the INSIM-FT-3D parameters are shown in Fig. 4.14 and the posterior (history-matched)

results are shown in Fig. 4.15. In Figs. 4.14 and 4.15 and similar figures, the vertical black line represent a separator between the historical period and the prediction period. As shown in Fig. 4.15, the history match of the oil rate of well P3 is the worst but the calculated oil rates from the ensemble of history-matched models bound the oil rate predicted from the true model during the 900 day "future" prediction period subsequent to the "historical" period. Moreover, the uncertainty in predictions is much lower than the uncertainty in predictions made from the ensemble of prior models. The prior and posterior relative permeabilities, respectively, are shown in Figs. 4.16a and 4.16b. It shows that the history-matched INSIM-FT-3D models reasonably bound the true relative permeabilities with a narrower uncertainty range compared with that of the prior models.



(a) layer 1 (b) layer 2 (c) layer 3

(d) layer 4 (e) layer 5 (f) layer 6

Figure 4.10: Horizontal absolute permeability map for a six-layer channelized reservoir. The dark blue areas indicate the shale zones; the light blue zones represent the levee facies and the yellow zones represent the channel facies. The color bar represents the scale of absolute permeability.

4.4 Example 3: Field Case

This example pertains to a reservoir with a strong bottom water drive which was

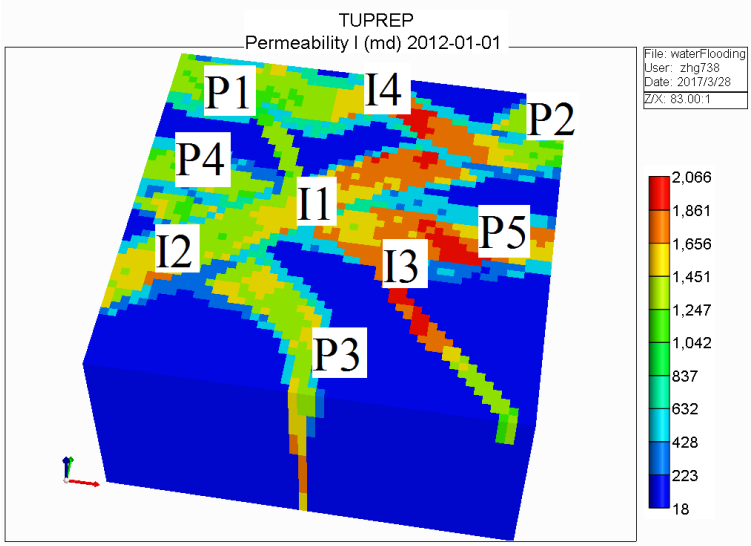


Figure 4.11: Well locations for the channelized reservoir.

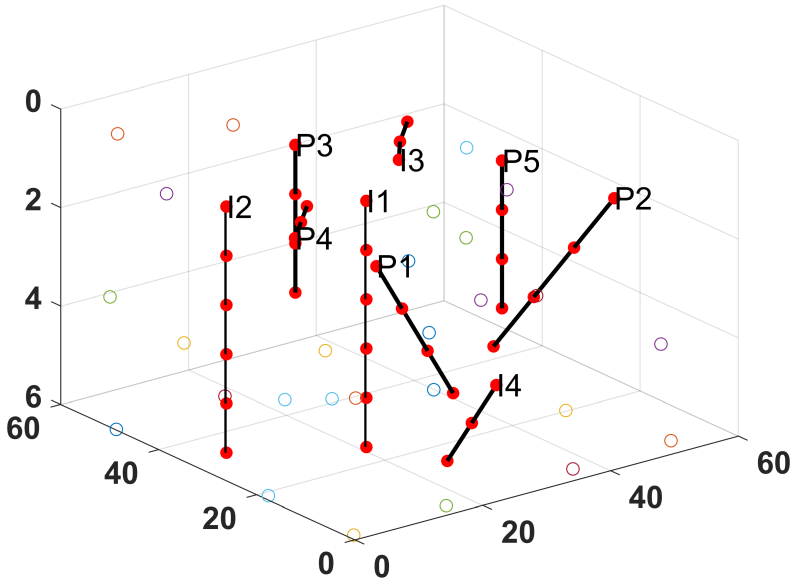


Figure 4.12: The well trajectories and actual and imaginary well nodes, channelized example.

tested in Chapter 2 to show that INSIM-FT can be applied for history matching field data. Due to the effects of the aquifer, the assumption of no gravity effects when applying INSIM-FT is not strictly correct. Here, we redo this example to investigate if a better history match and prediction can be obtained by applying the new INSIM-FT-3D data-driven model. The reservoir has 13 producers and no water injector exists; bottom-water drive is the major drive mechanism. The connection map can be seen in Fig. 2.24 with 26 imaginary wells

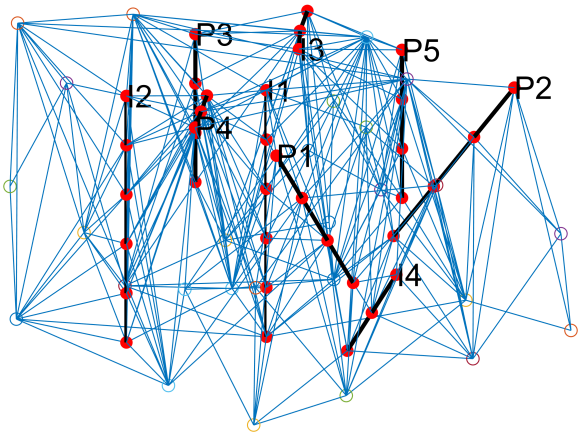


Figure 4.13: The connection map generated by Delaunay triangulation with modification, channelized example.

Table 4.1: Property of the channelized reservoir.

Property	Values
Oil compressibility, psi^{-1}	4×10^{-5}
Water compressibility, psi^{-1}	3.744×10^{-6}
Rock compressibility, psi^{-1}	6.103×10^{-5}
Oil viscosity, cp	2.2
Water viscosity, cp	1
Oil density, lbm/ft^3	56.93
Water density,lbm/ft^3	62.428

added to introduce more flow paths. In Fig. 2.24, the aquifer is represented by a single node that is placed in the depth below the datum depth of the water-oil contact. All the producers are operated under specified total liquid production rates that vary with time. The other reservoir properties including the generation of the initial ensemble of models is the same one given in Chapter 2. The ensemble size for ES-MDA is 300. The first 800 days of production data are history-matched and the subsequent 200 days are used for predictions. The oil rates of four representative wells calculated from the prior ensemble of INSIM-FT-3D models are shown in Fig. 4.17. The oil rates obtained from the history-matched INSIM-FT-3D models are compared side-by-side with the results from the history-

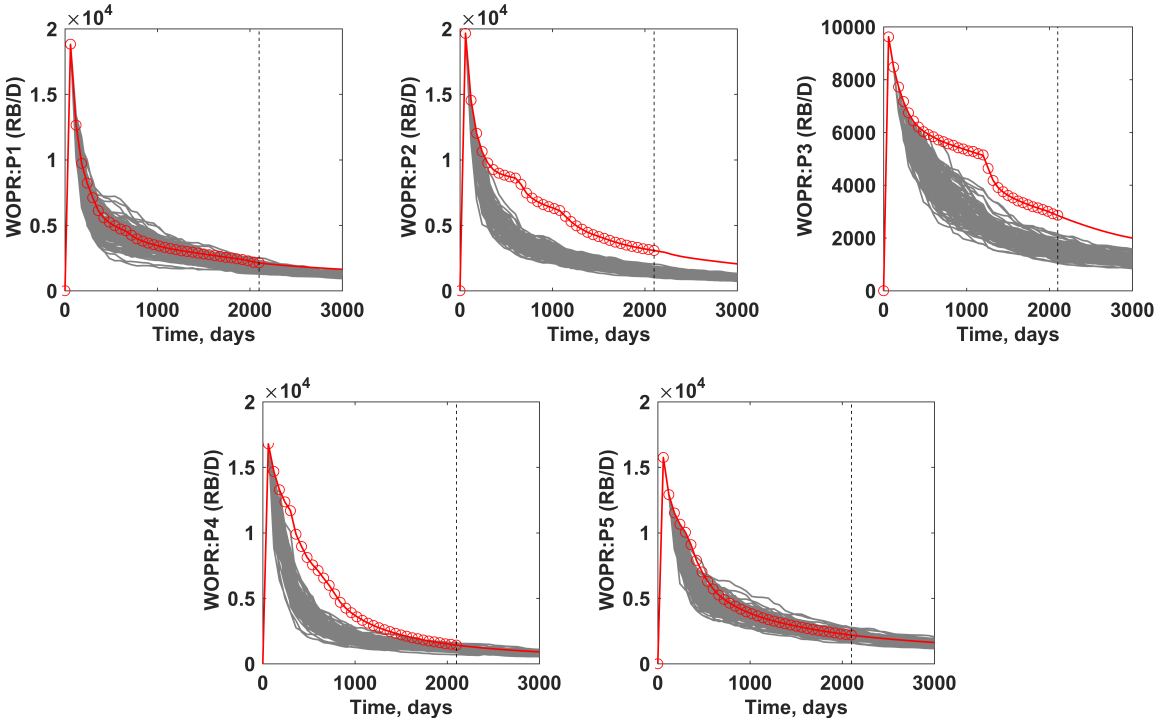


Figure 4.14: The estimated oil production rates obtained from the prior INSIM-FT-3D models. Red circles denote the observed oil production rates; red curves denote the true values of the oil production rate generated with CMG; gray curves represent the estimated oil rates obtained by running the 200 prior INSIM-FT-3D models.

matched INSIM-FT models without considering gravity (using the same connection map of Fig. 2.24) in Figs. 4.18 and 4.19. We can see that the history-matched INSIM-FT-3D models generates narrower uncertainty ranges for predicted oil production rates compared with that from the history-matched INSIM-FT models. For well W11, the INSIM-FT-3D history-matching models generates slightly better history-matches and predictions than those obtained with the original INSIM-FT history-matched models. The matching qualities for relative permeabilities are similar for the two methods and therefore are not shown here.

4.5 Example 4: Brugge Reservoir

The Brugge field is a synthetic reservoir developed by TNO as a benchmark case to test different methods for closed-loop reservoir management. A single realization of the Brugge reservoir example is used here to test the performance of INSIM-FT for handling a field-scale 3D reservoir with multiple wells. The top structure of the reservoir is shown in

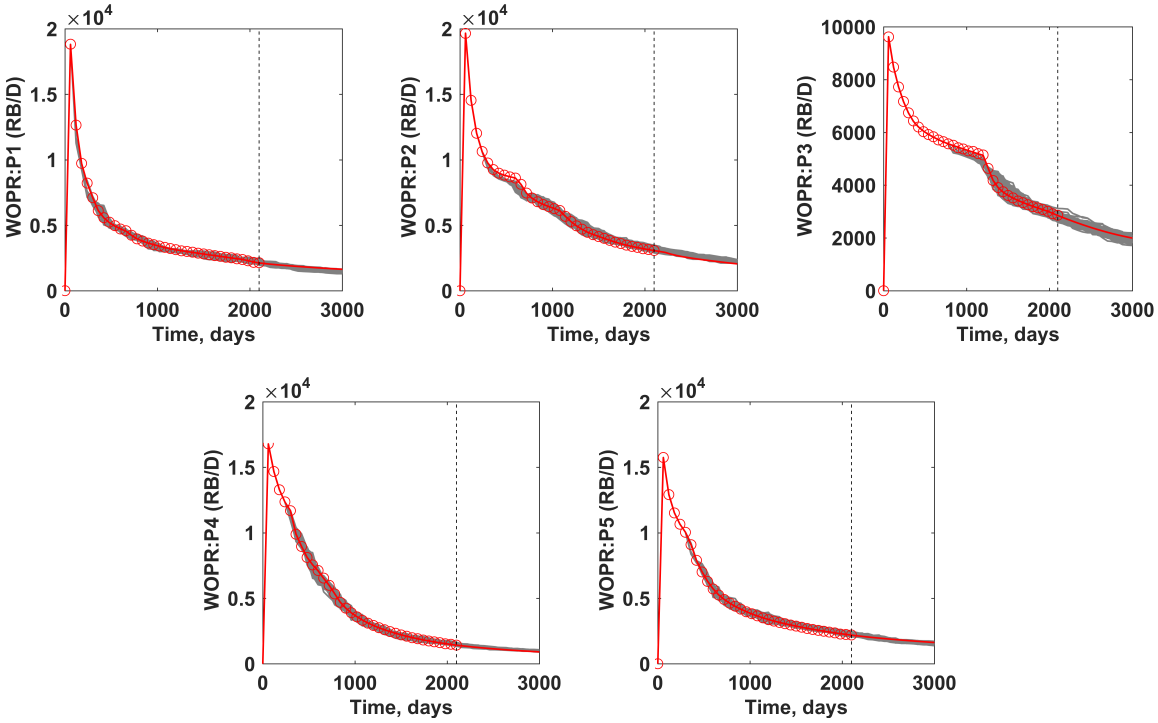
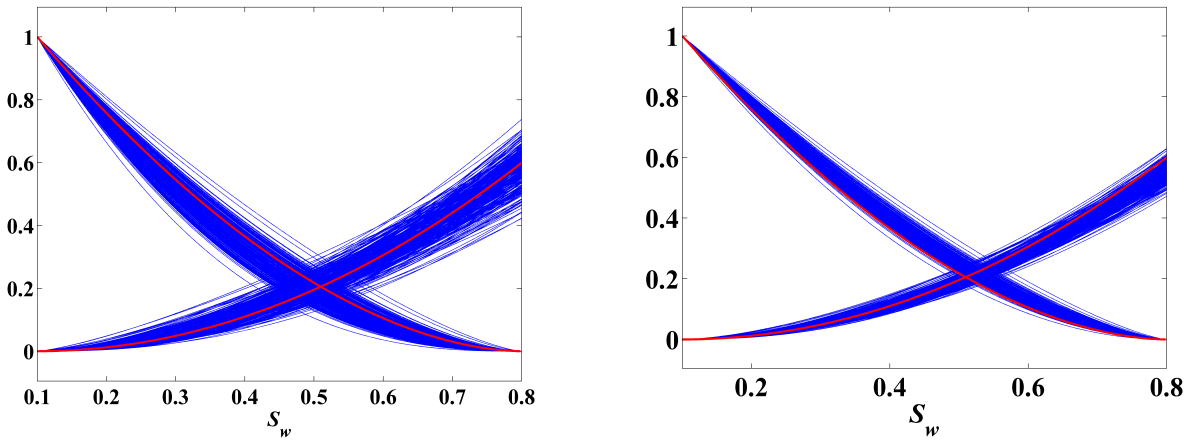


Figure 4.15: The estimated oil production rates obtained from the history-matched INSIM-FT-3D models. Red circles denote the observed oil production rates; red curves denote the true values of the oil production rate generated with CMG; gray lines represent the estimated oil rates obtained by running the 200 history-matched INSIM-FT-3D models.



(a) Prior relative permeabilities

(b) Posterior relative permeabilities

Figure 4.16: The prior and posterior relative permeabilities obtained with INSIM-FT-3D based on an ensemble size of 200. The blue curves represent the estimated relative permeabilities and the red curves represent the true relative permeabilities.

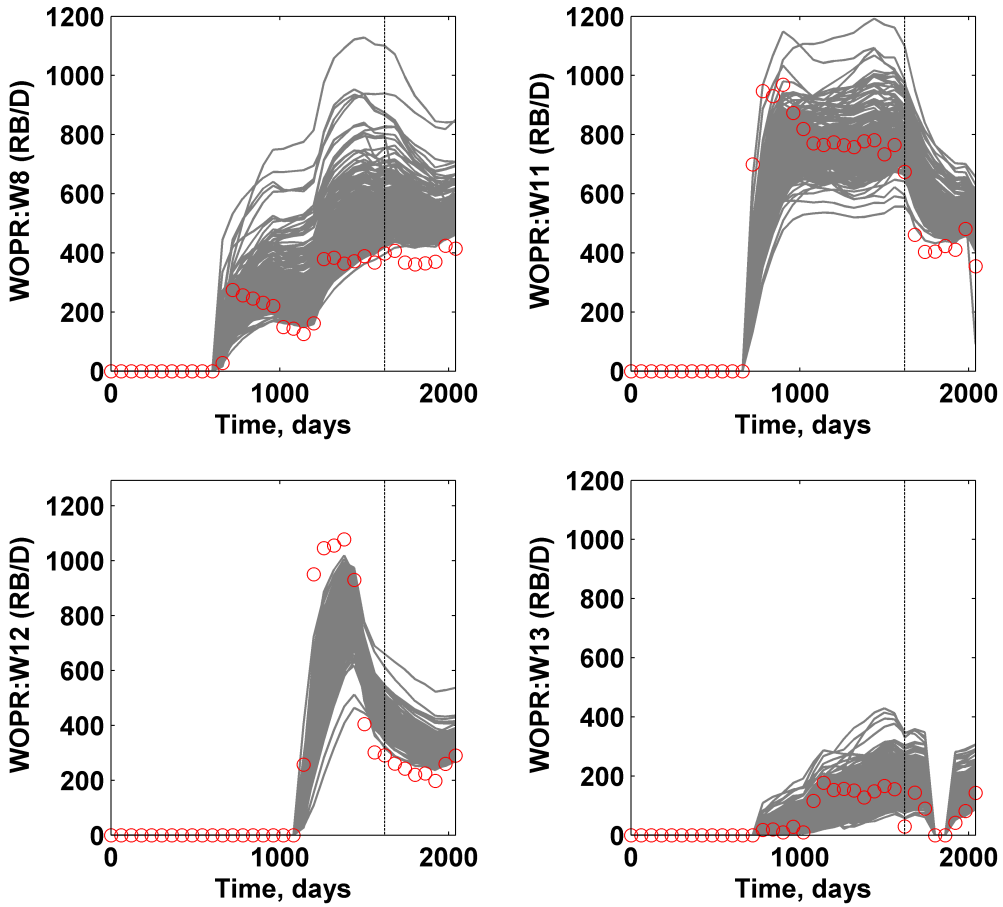
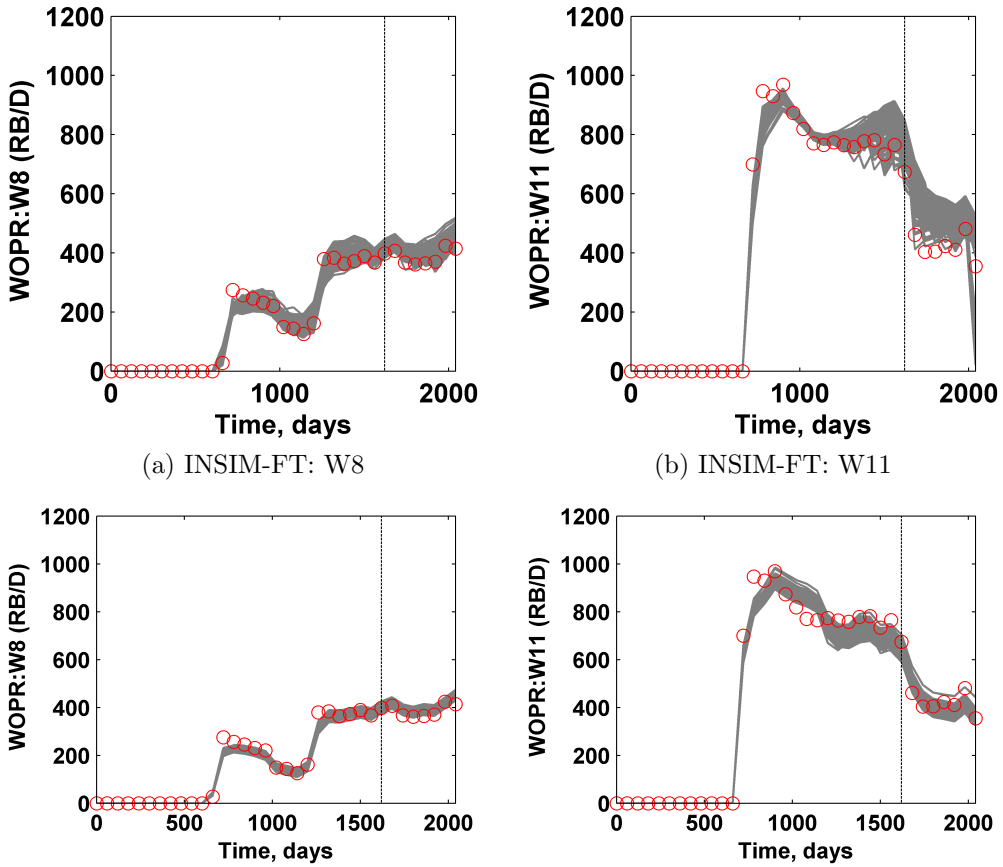


Figure 4.17: Prior oil production rates of the field example. Red circles are the observed oil production rates; gray curves represent the prior responses of the oil production rates obtained from the prior INSIM-FT models.

Fig. 4.20 (Chen, 2017), where 20 vertical production wells and 10 vertical injection wells are drilled. The true Eclipse model consists of four geological layers and nine reservoir simulation layers where each simulation layer has a 139×48 grid. Basic information on reservoir properties of the Brugge reservoir model can be found in Peters et al. (2010). All the wells are perforated in more than one simulation layer and the total number of perforations is 232. To apply INSIM-FT-3D, 200 imaginary wells are generated using Mitchell's best-candidate algorithm as shown in Fig. 4.21 and the resulting connection map generated with Delaunay triangulation is shown in Fig. 4.22. The production history of the Brugge reservoir is 30 years, but we only use the first ten years' production data generated with the true Eclipse model for history matching and prediction. Specifically, the observed data from the first 2,400



(a) INSIM-FT: W8

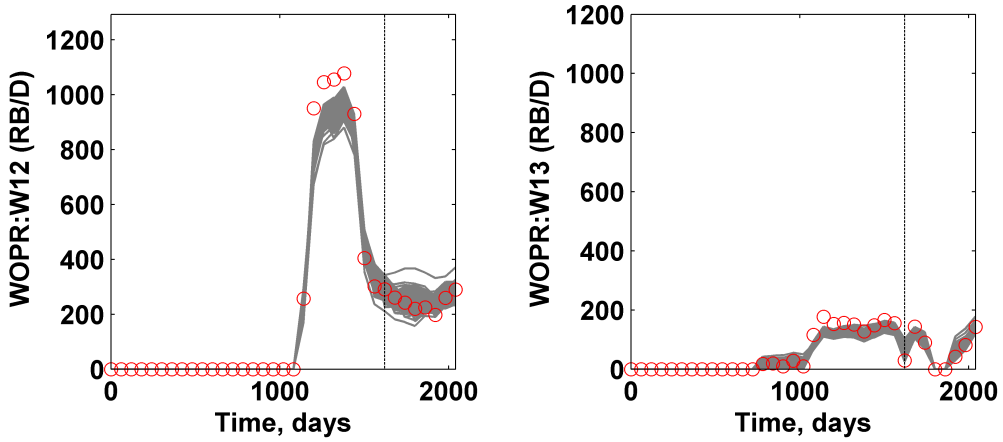
(b) INSIM-FT: W11

(c) INSIM-FT-3D: W8

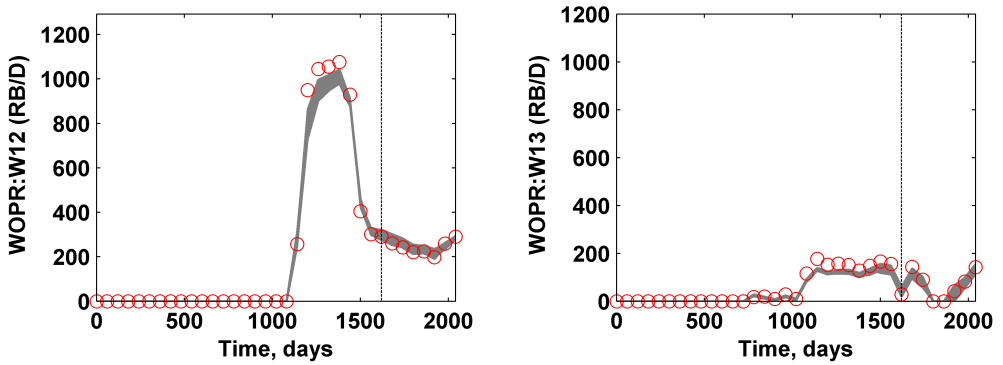
(d) INSIM-FT-3D: W11

Figure 4.18: History-matched oil production rates of the field example. Red circles are the observed oil rates; gray curves are the oil rates estimated with the history-matched INSIM-FT or INSIM-FT-3D models. The top two sub-figures are results obtained with INSIM-FT and the bottom two are obtained with INSIM-FT-3D.

days are for history matching and the following 1,200 days are used for future predictions. The true total rates for injectors and producers over the first ten years are obtained by running the Eclipse true model and are specified as the well operating conditions for history matching and future predictions with INSIM-FT-3D. The observed oil production rate data that are history matched are generated by adding uncorrelated Gaussian noise to the oil rates obtained with the true Eclipse model, where the standard deviation of the Gaussian noise is 5% of the true data. The ensemble size for ES-MDA is 200 and the initial ensemble is generated in the same way as described in the second example. ES-MDA with eight data assimilation steps is used to perform history-matching.



(a) INSIM-FT: W12 (b) INSIM-FT: W13



(c) INSIM-FT-3D: W12 (d) INSIM-FT-3D: W13

Figure 4.19: History-matched oil production rates of the field example. Red circles are the observed oil rates; gray curves are the oil rates estimated with the history-matched INSIM-FT or INSIM-FT-3D models. The top two sub figures are results obtained with INSIM-FT and the bottom two are obtained with INSIM-FT-3D.

The oil production rates calculated with the prior ensemble for 9 representative producers are shown in Fig. 4.23. The posterior oil production rates for the same 9 producers are shown in Fig. 4.24. Overall, the matching and prediction performance of INSIM-FT-3D is very good; the history-matched models bound the oil rate predicted from the true model during the prediction period. Fig. 4.25 shows the history-matched field oil production rates calculated with the 200 posterior INSIM-FT-3D realizations. The true field oil production rate curve is well bounded by the posterior results in both the history matching period and future prediction period.

To complete the discussion of the Brugge example, we need to make some critically

important comments. When history-matching dynamic data with ES-MDA, or any ensemble-based method, if the initial ensemble gives extremely biased predictions, it may require some modification in the application of ES-MDA to obtain an acceptable history-match. One possible fix is to modify the initial set of realizations so that the initial predictions are less biased. Here, however, a simpler option is applied. Applying the ES-MDA to the initial ensemble of realizations of the vector of model parameters failed to give an acceptable match of some wells. Thus, we identify the posterior model that gave the best overall data match using ES-MDA and define a new prior model of parameters by replacing the original prior mean by this specific posterior model of the vector of model parameters. Although the means are changed in the "new" prior, the prior covariance matrix is not changed. Using these updated normal distributions for the model parameters, a new initial ensemble of 200 models is generated and history matched with ES-MDA. The matches and predictions shown in Figs. 4.24 and 4.25 pertain to this second history-match. By way of illustration, Fig. 4.26 compares the history-match and predictions obtained for two production wells, P2 and P3, from the first and second history-match runs. Apparently, the performance of history matching and prediction is enhanced after the second round of history match. It is important to note that in all other examples shown in this work, only one ES-MDA history-matching run is performed.

One major advantage of the INSIM-FT-3D methodology presented here is that it does not require a detailed geological model to generate an ensemble of initial models for an ensemble-based history-matching process. A second major advantage is its computational efficiency. For the true reservoir model used in this example, a single run to simulate the production performance for the entire production history of 3,600 days, requires approximately 1.5 minutes. For the history matching period of the first 2,400 days, a single simulation run requires approximately one minute. If we applied ES-MDA with eight data assimilation steps and used an ensemble of 200 full-scale reservoir simulation models models, then it would require $8 \times 200 \times 1 = 1,600$ minutes to match the data for the 2,400 day historical period and then predict the reservoir performances for another 1,200 days. On the other hand, the

history-matching and predictions with INSIM-FT-3D required 60 minutes of total computational time. This reported time for INSIM-FT-3D includes both of the history-matching and prediction runs.

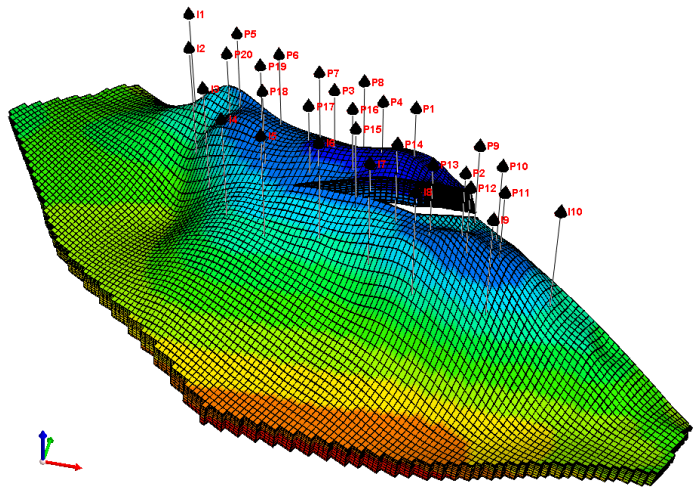


Figure 4.20: Top structure of Brugge field.

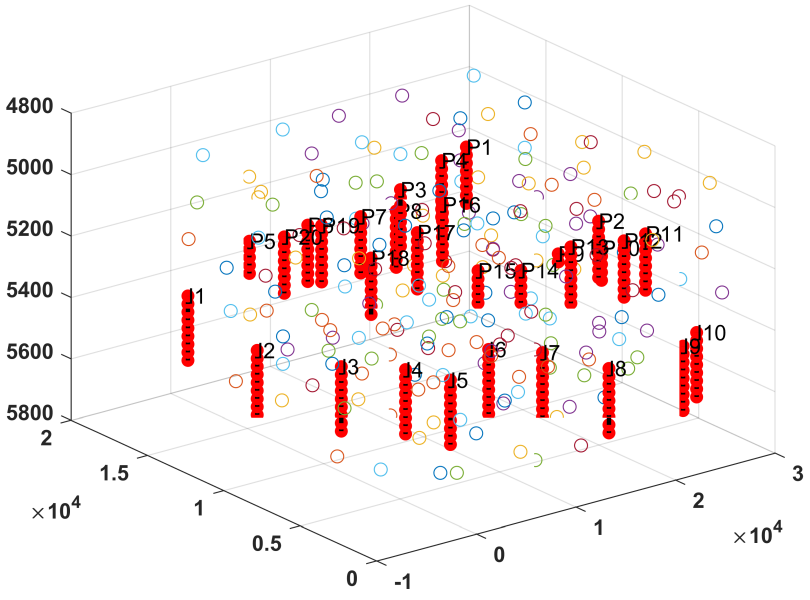


Figure 4.21: Well trajectories and actual and imaginary well nodes, Brugge example.

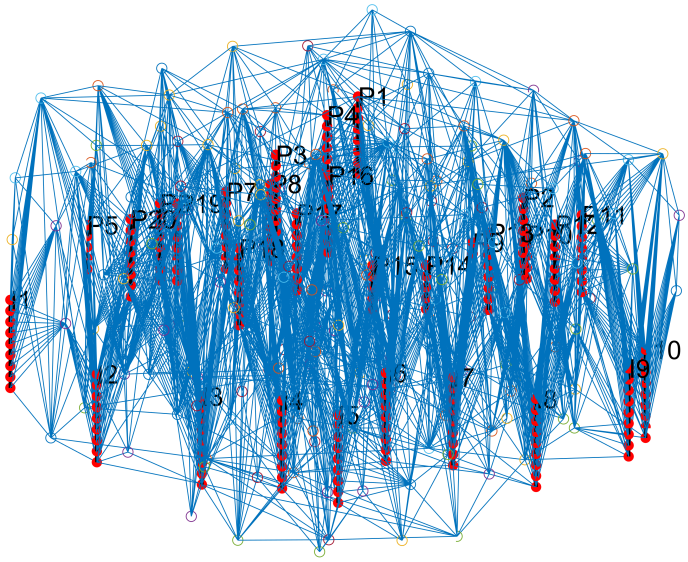
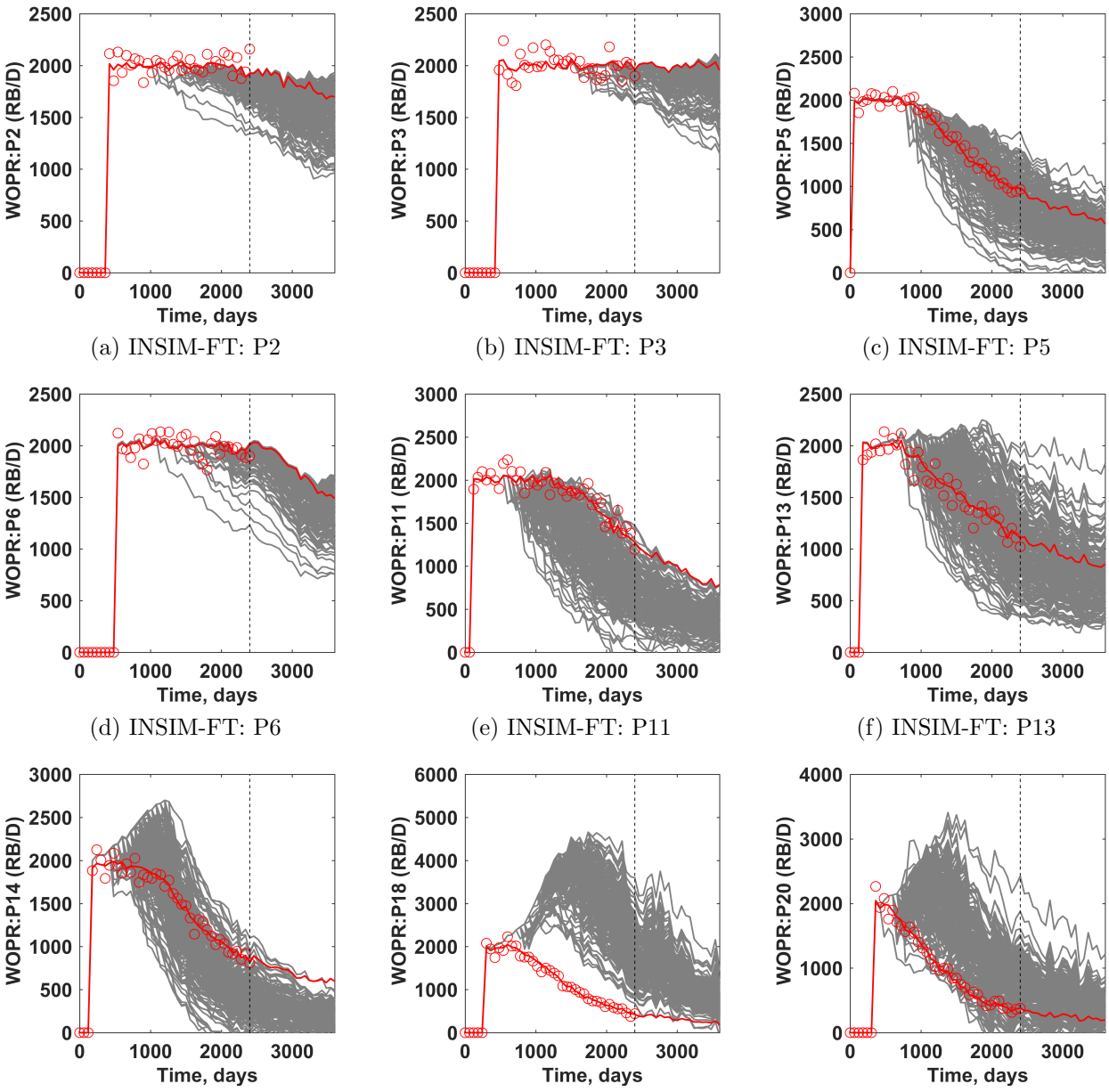


Figure 4.22: Connection map generated with Delaunay triangulation, Brugge example.

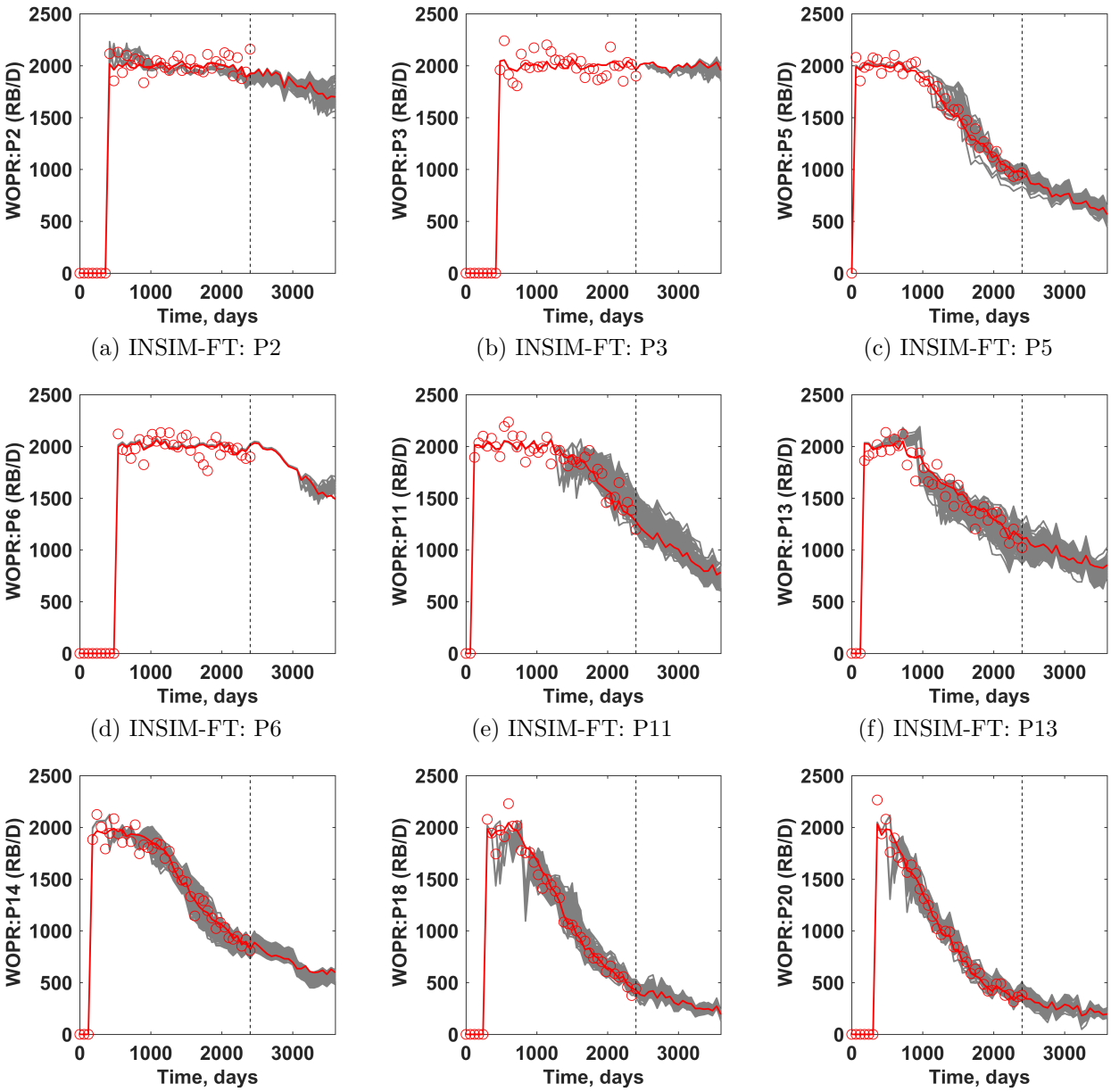


(a) INSIM-FT: P2 (b) INSIM-FT: P3 (c) INSIM-FT: P5

(d) INSIM-FT: P6 (e) INSIM-FT: P11 (f) INSIM-FT: P13

(g) INSIM-FT: P14 (h) INSIM-FT: P18 (i) INSIM-FT: P20

Figure 4.23: Prior oil production rates of the field example. Red circles are observed oil rates; gray curves are the oil rates estimated with INSIM-FT-3D models.



(a) INSIM-FT: P2 (b) INSIM-FT: P3 (c) INSIM-FT: P5

(d) INSIM-FT: P6 (e) INSIM-FT: P11 (f) INSIM-FT: P13

(g) INSIM-FT: P14 (h) INSIM-FT: P18 (i) INSIM-FT: P20

Figure 4.24: Posterior oil production rates of the Brugge example. Red circles are the observed oil rates; gray curves are the oil rates estimated with INSIM-FT-3D models.

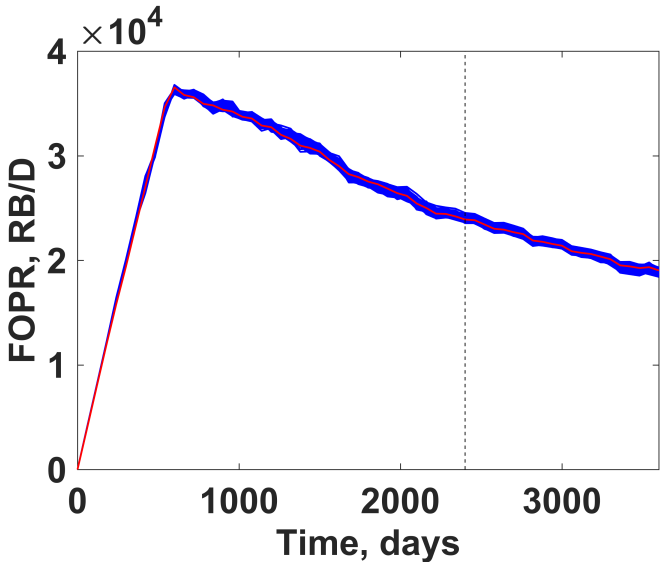
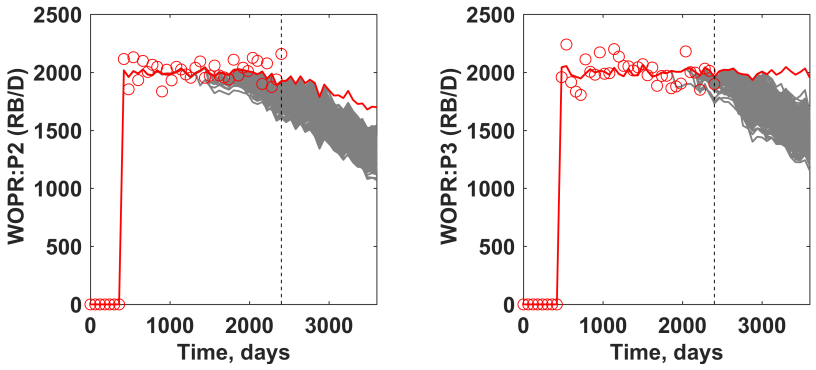


Figure 4.25: Posterior field oil production rates of the Brugge example. Red curve represents the true field oil production rates generated from the true Eclipse model; blue curves represent the field oil production rates calculated with the INSIM-FT-3D posterior realizations.



(a) First: P2 (b) First: P3

(c) Second: P2 (d) Second: P3

Figure 4.26: Comparison of history match and predictions for two production wells from the first and second history-match runs. The top two figures show the results for the first round of history-match and the bottom two figures show the results for the second round of history-match.

CHAPTER 5

PRODUCTION OPTIMIZATION WITH INSIM-FT-3D

In Chapter 4, we developed INSIM-FT-3D as a data-driven model for history matching and reservoir prediction. Compared to INSIM-FT, INSIM-FT-3D considers gravitational effects and allows wells with arbitrary well trajectories. Furthermore, an automatic workflow is provided to add imaginary wells and generate the connection map for history matching with INSIM-FT-3D. In this chapter, we investigate the possibility of using the history-matched INSIM-FT-3D model as the forward model when estimating the optimal well controls that maximize the net present value (NPV) for a water flood. The optimal NPV obtained with INSIM-FT are validated by inputting the optimal controls of INSIM-FT into the true model represented by a commercial reservoir simulation model. Also, we compare the optimization performance obtained by running INSIM-FT with the performance obtained by running the true model. To demonstrate the applicability of INSIM-FT-3D in production optimization for large-scale three-dimensional problems, two three-dimensional reservoirs are considered. The first one is a multi-layered channelized reservoir and the second one is the Brugge field case. Both the examples were used for history matching with INSIM-FT-3D in Chapter 4.

5.1 Formulations

The NPV function for a reservoir under waterflooding is defined in Eq. 3.1 which is repeated here as the following equation.

$$J(u) = \sum_{n=1}^{N_t} \left\{ \frac{\Delta t_n}{(1+b)^{\frac{t_n}{365}}} \Big[\sum_{j=1}^{P} (r_o \cdot q_{o,j}^n - c_w \cdot q_{w,j}^n) - \sum_{j=1}^{I} (c_{wi} \cdot q_{wi,j}^n) \Big] \right\}, \tag{5.1}$$

where u is vector of control variables for all wells and $J(u)$ represents the NPV value obtained with the well controls u.

To apply INSIM-FT-3D for maximizing the NPV function, we assume that the target reservoir has been produced for a long enough time so that abundant production data can be collected for characterizing the reservoir by history matching with INSIM-FT-3D. Then, the history-matched INSIM-FT-3D model is treated as the forward model to optimize the well controls in the remaining life-time of the reservoir in order to obtain a maximum NPV value. Note that, we use ES-MDA for history matching and obtain multiple history-matched INSIM-FT-3D models, i.e., multiple realizations of the parameters defining an INSIM-FT-3D model. As shown in Chapter 4, the vector of parameters, m, that define a INSIM-FT-3D model include the transmissibilities, pore volumes, relative permeabilities and well indices for wells with multi-segments. The history-matched INSIM-FT-3D model with the best data match is denoted by m_{best} and m_{best} will be used as the forward model to perform production optimization.

The production optimization period is equally divided into N_c consecutive time intervals with each part referred to as a control step. On each control step, a constant well operating condition, bottom-hole pressure or total flow rate, is specified for each well. The total set of these rates and pressures represent the well controls. By optimizing the operating conditions of each control step for each well, one can obtain an optimal NPV value. Determining the well controls that maximize the NPV of production over the remaining life of the reservoir is the optimal well control problem. However, the original INSIM-FT-3D model in Eq. 4.4 does not allow BHPs as well operating conditions. In order to optimize NPV with INSIM-FT-3D by tuning BHPs, the original INSIM-FT-3D formulation is discussed in the remainder of this section.

For wells with a single perforation, the pressure equation of the original INSIM-FT-3D model is given by Eq. 4.4 which is repeated in the following equation.

$$\sum_{j=1}^{n_{c,i}} T_{i,j}^{n-1}\Big(p_j^n - p_i^n - \big[(\gamma_{o,i,j}^{n-1}\frac{\lambda_o(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})} + \gamma_{w,i,j}^{n-1}\frac{\lambda_w(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})})(D_j - D_i)\big]\Big) + q_{t,i}^n$$
$$= \frac{1}{5.615}\frac{c_{t,i}^{n-1}V_{p,i}^{n-1}}{\Delta t_n}(p_i^n - p_i^{n-1}). \quad (5.2)$$

To allow pressure control for production optimization, the source/sink term, $q_{t,i}^n$ in Eq. 5.2, is replaced by

$$q_{t,i}^n = -\mathrm{WI}_i\lambda_{t,i}^{n-1}(p_i^n - p_{wf,i}^n). \quad (5.3)$$

Then, in the production optimization period, the pressure equation for INSIM-FT-3D with BHP control is given by

$$\sum_{j=1}^{n_{c,i}} T_{i,j}^{n-1}\Big(p_j^n - p_i^n - \big[(\gamma_{o,i,j}^{n-1}\frac{\lambda_o(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})} + \gamma_{w,i,j}^{n-1}\frac{\lambda_w(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{n-1})})(D_j - D_i)\big]\Big)$$
$$- \mathrm{WI}_i\lambda_{t,i}^{n-1}(p_i^n - p_{wf,i}^n) = \frac{1}{5.615}\frac{c_{t,i}^{n-1}V_{p,i}^{n-1}}{\Delta t_n}(p_i^n - p_i^{n-1}). \quad (5.4)$$

By applying the modification of Eq. 5.4, BHPs can be specified as well controls in INSIM-FT-3D for wells with a single perforation. Using the same procedure described in Chapter 3, the well indices in Eq. 3.9 are determined using the information predicted from the history-match run corresponding to m_{best}. Specifically, the well index for well i that only has one perforation is given by

$$\overline{\mathrm{WI}}_i = \frac{1}{n_t}\sum_{k=1}^{n_t}\frac{-q_{t,i}^k}{(p_i^k - p_{wf,\mathrm{obs},i}^k)\lambda_{t,i}^k}, \quad (5.5)$$

where the superscript k represents the INSIM-FT-3D step; n_t is the number of INSIM-FT-3D steps for the history-matching period; $p_{wf,\mathrm{obs},i}$ is the observed BHP of well i; $q_{t,i}$ is the specified total well rate of well i; p_i is the node pressure of volume i and $\lambda_{t,i}$ is the total mobility for node i; both p_i and $\lambda_{t,i}$ are predicted from the INSIM-FT-3D history-match run

based on m_{best}. The water saturation equation for INSIM-FT-3D is defined by Eq. 4.19 for the connection between (i, j),

$$\frac{\partial S_w(x,t)}{\partial t} + \frac{5.615 q^n_{t,i,j} L_{i,j}}{V^n_{p,i,j}} \frac{\partial f_w(x,t)}{\partial x} = 0 \quad \text{for} \quad 0 \le x \le L_{i,j}, \quad t^{n-1} \le t \le t^n. \tag{5.6}$$

After solving the INSIM-FT-3D model for pressure and water saturation, the total flow rate at well i with a single perforation that is controlled by BHP can be calculated by

$$q^n_{t,i} = -\overline{\text{WI}}_i \lambda^{n-1}_{t,i} (p^n_i - p^n_{wf,i}), \tag{5.7}$$

were a negative value of $q^n_{t,i}$ represents well i is a producer and a positive value represents well i is an injector. If well i is a producer, the oil production rate of well i at time level n is computed by

$$q^n_{o,i} = \overline{\text{WI}}_i \lambda^{n-1}_{t,i} (p^n_i - p^n_{wf,i}) \cdot (1 - f_w(S^n_{w,i})), \tag{5.8}$$

and the water production rate of well i is given by

$$q^n_{w,i} = \overline{\text{WI}}_i \lambda^{n-1}_{t,i} (p^n_i - p^n_{wf,i}) \cdot f_w(S^n_{w,i}), \tag{5.9}$$

where $f_w(S_w)$ is the fractional flow function with gravitational effects which is defined in Eq. 4.21.

For wells with multiple perforations, the original pressure equation for well w is defined for each perforation-controlled volume i_k, for $k = 1, 2, \ldots n^w_{wc}$, which is given by combining Eqs. 4.10 and 4.11,

$$\sum_{j=1}^{n_{c,i_k}} T^{n-1}_{i_k,j} \Big(p^n_j - p^n_{i_k} - \big[(\gamma^{n-1}_{o,i_k,j} \frac{\lambda_o(S^{n-1}_{w,i_k,j})}{\lambda_t(S^{n-1}_{w,i_k,j})} + \gamma^{n-1}_{w,i_k,j} \frac{\lambda_w(S^{n-1}_{w,i_k,j})}{\lambda_t(S^{n-1}_{w,i_k,j})}) (D_j - D_{i_k}) \big] \Big)$$
$$+ \text{WI}^w_{i_k} \lambda_t(S^{n-1}_{w,i_k}) (p^n_{\text{wf}} + H^{n-1}_{w,k} - p^n_{i_k}) = \frac{1}{5.615} \frac{c^{n-1}_{t,i_k} V^{n-1}_{p,i_k}}{\Delta t_n} (p^n_{i_k} - p^{n-1}_{i_k}). \tag{5.10}$$

The fact that the total flow rate from well w must be equal to the sum of rates through all

the perforated segments yields

$$q_{t,w}^n = \sum_{k=1}^{n_{wc}^w} \mathrm{WI}_{i_k}^w \lambda_t(S_{w,i_k}^{n-1})(p_{\mathrm{wf}}^n + H_{w,k}^{n-1} - p_{i_k}^n), \tag{5.11}$$

where $H_{w,k}$ is the well-bore pressure difference between the center of the kth segment of well w and the bottom-hole datum depth, which is defined by Eq. 4.12. By specifying the total liquid rate of well w, $q_{t,w}$, at each INSIM-FT-3D step, the reservoir pressure associated with each perforation and the BHP of well w are obtained simultaneously by solving Eq. 5.10 and 5.11. Ideally, we could specify the BHP of well w in 5.10 at each INSIM-FT-3D time step during the optimization period to run the INSIM-FT-3D forward model so that the NPV value given by Eq. 3.1 can be estimated. However, since the well indices of perforations for well w, $\mathrm{WI}_{i_k}^w$, are obtained by only matching the historical data of oil production rate, the well indices are not conditioned to the BHP data of well w. Therefore, the NPV value estimated with the well indices is not reliable. Here, we provide a procedure to estimate $\widehat{\mathrm{WI}}_{i_k}^w$'s for production optimization, where $\widehat{\mathrm{WI}}_{i_k}^w$ represents the adjusted well index of kth perforation of well w by incorporating the BHP data.

First, we run the INSIM-FT-3D forward model with m_{best} and obtain the node pressures (p_{i_k}), saturations (S_{w,i_k}) and total liquid rates (q_{t,i_k}) for each perforation of well w during the history-matching period. Then, the well indices of well w for optimizing the production of the remaining life time of the reservoir are estimated by

$$\widehat{\mathrm{WI}}_{i_k}^w = \frac{q_{t,i_k}^n}{\lambda_t(S_{w,i_k}^{n-1})(p_{\mathrm{wf,obs}}^n + H_{w,k}^{n-1} - p_{i_k}^n)}, \tag{5.12}$$

where $p_{\mathrm{wf,obs}}$ represents the observed BHPs of well w. One difficulty with this procedure is that due to the approximate nature of INSIM-FT-3D, a $\widehat{\mathrm{WI}}_{i_k}^w$ calculated from Eq. 5.12 varies with time, whereas, well indices are by definition independent of time. Thus instead

of solving Eq. 5.12, for $\widehat{\mathrm{WI}}_{i_k}^{w}$, we use an average over time which is given by

$$\overline{\widehat{\mathrm{WI}}}_{i_k}^{w} = \frac{1}{n_t}\sum_{n=1}^{n_t}\frac{q_{t,i_k}^{n}}{\lambda_t(S_{w,i_k}^{n-1})(p_{\mathrm{wf,obs}}^{n}+H_{w,k}^{n-1}-p_{i_k}^{n})}, \tag{5.13}$$

where n_t is the number of time steps used for the history-match run with INSIM-FT-3D. With the well indices given by Eq. 5.13 specified in the production optimization period, we can now specify BHPs as well controls for optimization. The total liquid rate of well w at time level n with BHP controls is calculated with Eq. 5.14, which is given by

$$q_{t,w}^{n} = \sum_{k=1}^{n_{wc}^{w}}\left[\overline{\widehat{\mathrm{WI}}}_{i_k}^{w}\lambda_t(S_{w,i_k}^{n-1})(p_{\mathrm{wf}}^{n}+H_{w,k}^{n-1}-p_{i_k}^{n})\right]. \tag{5.14}$$

If well w represents a production well, the total oil and water rate for well w at time level n at the optimization period with BHP controls, $q_{o,w}^{n}$ and $q_{w,w}^{n}$, respectively, are given by

$$q_{o,w}^{n} = \sum_{k=1}^{n_{wc}^{w}}\left[\overline{\widehat{\mathrm{WI}}}_{i_k}^{w}\lambda_o(S_{w,i_k}^{n-1})(p_{\mathrm{wf}}^{n}+H_{w,k}^{n-1}-p_{i_k}^{n})\right], \tag{5.15}$$

and

$$q_{w,w}^{n} = \sum_{k=1}^{n_{wc}^{w}}\left[\overline{\widehat{\mathrm{WI}}}_{i_k}^{w}\lambda_w(S_{w,i_k}^{n-1})(p_{\mathrm{wf}}^{n}+H_{w,k}^{n-1}-p_{i_k}^{n})\right]. \tag{5.16}$$

With the method to compute the phase rates at INSIM-FT-3D steps for different wells, the NPV value for a given set of well controls can be estimated by Eq. 5.1.

5.2 Optimization Procedure

The optimization method to maximize the NPV value for waterflooding considered here is steepest ascent with simplex gradient as implemented by Do and Reynolds (2013). To optimize with INSIM-FT-3D, the optimization period is divided into a sequence of N_c control steps. By tuning the well controls specified for each control step, NPV defined in Eq. 5.1 is maximized. Here, we consider injectors controlled by total injection rates and producers controlled by BHPs. The optimization procedure is similar as that described in the fourth

section of Chapter 3, expect that Chapter 3 used an EnOpt gradient (Chen et al., 2009) based on a slight modification by Do and Reynolds (2013) for optimization. Optimization is done in terms of the normalized control variables defined in Eq. 3.6. Then, for well ℓ and jth control step, given the value of the normalized control variable x_j^ℓ, the corresponding value of control variable, u_j^ℓ, is given by

$$u_j^\ell = x_j^\ell \cdot (u_j^{\ell,up} - u_j^{\ell,low}) + u_j^{\ell,low}, \tag{5.17}$$

where $u_j^{\ell,up}$ and $u_j^{\ell,low}$, respectively, represent the upper and lower bound of jth control variable for well ℓ .

With the vector of control variables for all wells, u, the NPV value of Eq. 5.1 can be estimated with the forward INSIM-FT-3D run based on m_{best}. Throughout this chapter, we let x represent the vector of normalized variables corresponding to the control variable u defined in Eq. 3.2. Here, the NPV value is maximized using steepest ascent algorithm based on a stochastic gradient. For kth iteration, the algorithm is written as

$$x^{k+1} = x^k + \alpha_k \left[\frac{d_k}{\|d_k\|_\infty} \right], \tag{5.18}$$

for $k = 0, 1, \ldots$ until convergence, where α_k is the step size with an initial value of 0.1; d_k is the search direction. Note that in Eq. 5.18, since x is normalized between [0,1], d_k is normalized by its infinity norm, which will ensure the largest change of the control variable for iteration k will be less than or equal to 0.1. In Chapter 3, we use a smoothed EnOpt gradient as the search direction based on the implementation of Do and Reynolds (2013) and the corresponding d_k is given by

$$d_{k,\text{EnOpt}} = \frac{C_x}{N_p} \sum_{j=1}^{N_p} (\hat{x}^j - x^k)[J(\hat{x}^j) - J(x^k)], \tag{5.19}$$

where $J(x)$ represents the NPV corresponding to the normalized well control x; $\hat{x}^j$ represents the perturbed well controls around x^k in order to calculate the approximate gradient of NPV

with respect to x^k and N_p is the number of perturbed well controls. To compute $d_{k,\text{EnOpt}}$ in Eq. 5.19, $\hat{x}^j$, $j = 1, 2, \ldots N_p$ are required to be randomly generated by sampling from the normal distribution of $\mathcal{N}(x^k, C_x)$, where C_x is the covariance matrix for generating the ensemble of well controls. To avoid abrupt changes in a well's control from one time step to another, C_x is designed to impose temporal correlation on controls at each well which is defined by

$$C_x = \begin{bmatrix} C_x^1 & 0 & . & 0 \\ 0 & C_x^2 & . & 0 \\ . & . & . & . \\ 0 & 0 & . & C_x^{N_w} \end{bmatrix}, \tag{5.20}$$

where N_w is the number of wells and C_x^ℓ is a well-control correlation matrix for well ℓ. The ith row and jth column of C_x^ℓ is calculated with a spherical model which is given by Eq. 3.17. Theoretically, the expectation of $\frac{1}{N_P}\sum_{j=1}^{N_p}(\hat{x}^j - x^k)[J(\hat{x}^j) - J(x^k)]$ is proved equal to $C_x \cdot \nabla J(u^k)$ (Do and Reynolds, 2013). Therefore, the search direction given by Eq. 5.19 has an expectation value of $C_x^2 \cdot \nabla J(u^k)$.

Note that sampling from $\mathcal{N}(x^k, C_x)$ does not ensure that all the generated well control variables are in the bound of [0,1]. In case that the sampled $\hat{x}^j$ value is not in [0,1], sometimes we need to implement truncation to enforce the bound constraints; otherwise, nonphysical well controls may be generated, e.g. negative injection rates. However, if the perturbed well controls are truncated, the expectation of $d_{k,\text{EnOpt}}$ of Eq. 5.19 is no longer equal to $C_x^2 \cdot \nabla J(u^k)$ and may cause the inaccuracy of the estimated gradient. To resolve this issue, we apply a simplex gradient (Do and Reynolds, 2013) without the requirement that the perturbed controls must satisfy $\mathcal{N}(x^k, C_x)$ in order to obtain the correct gradient approximation. To compute the simplex gradient, N_p sets of normalized controls are generated by sampling from $\mathcal{N}(x^k, C_x)$ as for EnOpt. Then, we implement truncation to ensure all the generated control variables are in the bound of [0,1]. Next, we can compute the simplex gradient of NPV based on each set of perturbed well controls. For jth perturbed well controls, the simplex gradient

is given by

$$d_{k,j} = \left((\hat{x}^j - x^k)^+\right)^T \left(J(\hat{x}^j) - J(x^k)\right), \tag{5.21}$$

where $\hat{x}^j$ represents the jth perturbed well controls which have been truncated for control variables outside the interval of [0,1]; the plus sign represents pseudo inverse, and the search direction used in steepest ascent of Eq. 5.18 is

$$d_{k,\text{simplex}} = C_x \frac{1}{N_p} \sum_{j}^{N_p} d_{k,j}. \tag{5.22}$$

Now, the expectation of $d_{k,\text{simplex}}$ can be proved equal to $C_x \nabla J(u^k)$ even if the perturbed controls are truncated.

Complete Workflow. The complete workflow for production optimization with INSIM-FT-3D is summarized below.

1. Specify the maximum number of total iterations, $N_{maxiter}$, allowed in steepest ascent and the maximum number of simulation runs, N_{maxsim}, allowed for optimization.

2. Choose the best history-matched INSIM-FT-3D model, m_{best} obtained using ES-MDA for production optimization. Compute the well indices for optimization using Eqs. 5.5 and 5.13 for each perforation of producers based on the history-match run according to m_{best}.

3. Start production optimization and set $k = 0$, where k is the iteration index for steepest ascent (Eq. 5.18). Use the observed BHP data and total injection and production rate data at the end of production history to estimate the initial guesses of controls for all control variables and generate the initial normalized control vector x^0.

   For $k = 0, 1, \ldots N_{maxiter}$,

   - Generate N_p perturbations of the normalized control vector, x^k, by sampling from

the Gaussian distribution $\mathcal{N}(x^k, C_x)$; the perturbations are given by

$$\hat{x}^j = x^k + C_x^{1/2} Z_j,\ j = 1, 2, \ldots N_p,$$

where

$$Z_j \sim \mathcal{N}(0, I).$$

If a component of $\hat{x}^j$ is outside the interval of [0,1], apply truncation to ensure all components of $\hat{x}^j$ satisfy the bound constraints.

- Run INSIM-FT-3D for each set of well controls, $\hat{u}^j = \hat{x}^j \left[u^{\text{up}} - u^{\text{low}}\right] + u^{\text{low}}$, $j = 1, 2, \ldots N_p$ to obtain corresponding NPV values.
- Compute the simplex gradient given by Eq. 5.22.
- Do linesearch based on the steepest ascent algorithm of Eq. 5.18 to obtain the updated control vector $x^{k+1} = x^k + \alpha_k \left[\frac{d_k}{\|d_k\|_\infty}\right]$. The initial stepsize for α_k is set to 0.1. If the linesearch can not find a control vector which increases the value of NPV in the maximum allowable number of stepsize cuts, then we set x^{k+1} to the set of controls obtained during the line search that gives the highest NPV. In the examples we consider, the maximum number of stepsize cuts is set to 5.
- If the number of maximum simulation runs N_{maxsim} or maximum iteration $N_{maxiter}$ are reached, or both of the following conditions are satisfied,

$$\frac{|J(x^{k+1}) - J(x^k)|}{\max[J(x^k), 1.0]} \leq \varepsilon_J, \tag{5.23}$$

$$\frac{\|x^{k+1} - x^k\|}{\max[\|x^k\|_2, 1.0]} \leq \varepsilon_x, \tag{5.24}$$

then terminate iteration. In the examples, we use $\varepsilon_J = 10^{-4}$ and $\varepsilon_x = 10^{-3}$.

End For

4. Record the values of u for x^k that gives the highest NPV as the optimal well controls.

5.3 Example 1: Multilayer Channelized Reservoir.

The first example pertains to a channelized reservoir with multi-layers as shown in Fig 4.10. The true model is represented by a CMG model with six layers. Each layer has a 50 by 50 grid and the size of each gridblock is 200 ft by 200 ft by 10 ft. The reservoir has four injectors and five producers with the well locations and trajectories shown in Figs. 4.11 and 4.12. More information on this reservoir is presented in the second example of Chapter 4. The history matching period lasts 1,500 days and the subsequent 1,500 days are for production optimization based on the best-matched INSIM-FT-3D model. The number of control steps, N_c, is set equal to 25, with each control step equal to 60 days in duration. The total number of control variables is 225, which is equal to N_c times the number of wells. During the optimization period, the well controls of injectors are water-injection-rate controls and BHP controls are used for producers. The well indices of producers for pressure controls are computed with Eqs. 5.5 and 5.13 from a single forward prediction run of INSIM-FT-3D with m_{best}. The upper and lower bounds, respectively, for all BHP controls are 1,000 and 3,800 psi, where 3,800 is the initial reservoir pressure; injection rate controls are constrained to be within [0, 15,000] RB/day. The initial guesses of the optimal controls are given by the controls at the end of history matching period, which are 2,500 psi for BHP of all producers and 10,000 RB/day for injection rates of all injectors.

The economic factors for optimization, see Eq. 5.1, are given by $r_o = 50$ USD/RB, $c_w = 5$ USD/RB, $c_{wi} = 2$ USD/RB and $b = 0.1$. For optimization with simplex gradient, the number of perturbed well controls for gradient estimate is 20. The maximum number of stepsize cuts for line search is 5; the initial search step for steepest ascent is 0.1; the time-correlation length to enforce the temporal smoothness of well controls is 420 days; the standard deviation of the spherical model we use to generate C_x defined in 3.17 is set to 0.01; the maximum number of allowable simulation runs is 1,000; the maximum number of allowed iterations is 60.

We perform production optimization based on running the history-matched INSIM-FT-3D model and the true CMG model as the forward model respectively. The NPV values versus forward runs obtained with the two different procedures are compared in Fig. 5.1. For this example, the total computational time spent on optimization with INSIM-FT-3D as the forward model is 10 minutes whereas it takes 11 hours to finish the optimization procedure using CMG as the forward model. Within the same number of iterations, INSIM-FT-3D yields an optimal NPV value, which is about 300 million USD more than the initial NPV estimated with INSIM-FT-3D. Similar NPV gain is achieved by optimization directly with the CMG true model, although the initial NPV estimated with CMG is 100 million dollars higher than is estimated with INSIM-FT-3D due to the differences between the two models. Figs. 5.2 and 5.3 show the comparisons of optimal well controls for INSIM-FT and CMG. Overall, the optimal well controls obtained with INSIM-FT-3D and CMG are quite similar for some wells, e.g., P2-P4. Both the results show that, in order to achieve the best production performance, P2 and P4 are expected to be shut in whereas P3 should be operated under the minimum allowable BHP.

We also compute the "true" NPV based on the optimal well controls generated with INSIM-FT-3D by inputting these optimal controls into the true CMG model; see the second row of Table 5.1. We can see that the NPV value obtained by using the INSIM-FT-3D optimal controls in the true model is 2,406 million USD, which is quite close to the value obtained with direct optimization with CMG (2,495 million USD). As shown in Table 5.1, the optimal NPV values obtained with three different scenarios are quite close to each other, which validates that our INSIM-FT-3D model can be used for generating a set of reliable optimal well controls for NPV maximization.

5.4 Example 2: Brugge Reservoir

The Brugge reservoir is a synthetic field designed as a benchmark case for closed-loop reservoir management Peters et al. (2010). Originally, the geological uncertainty of this reservoir is represented by 104 Eclipse reservoir simulation models. In Chapter 4, we

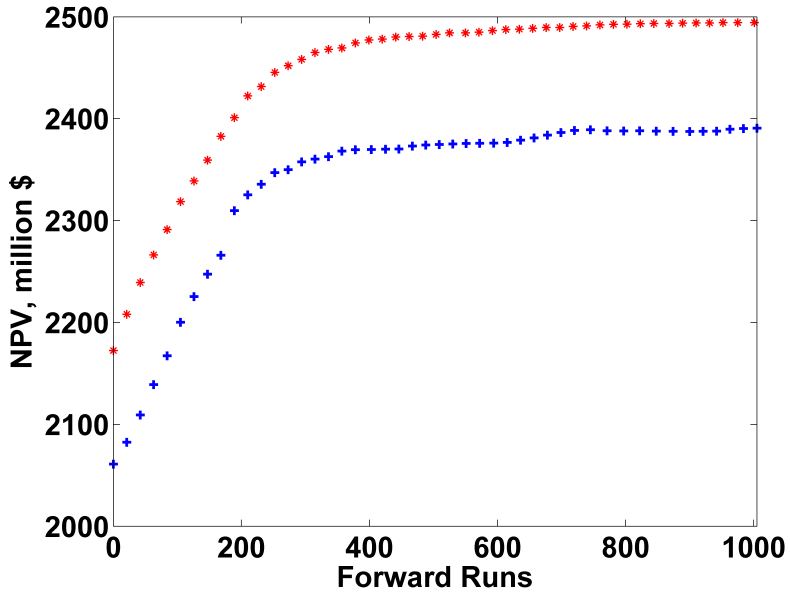
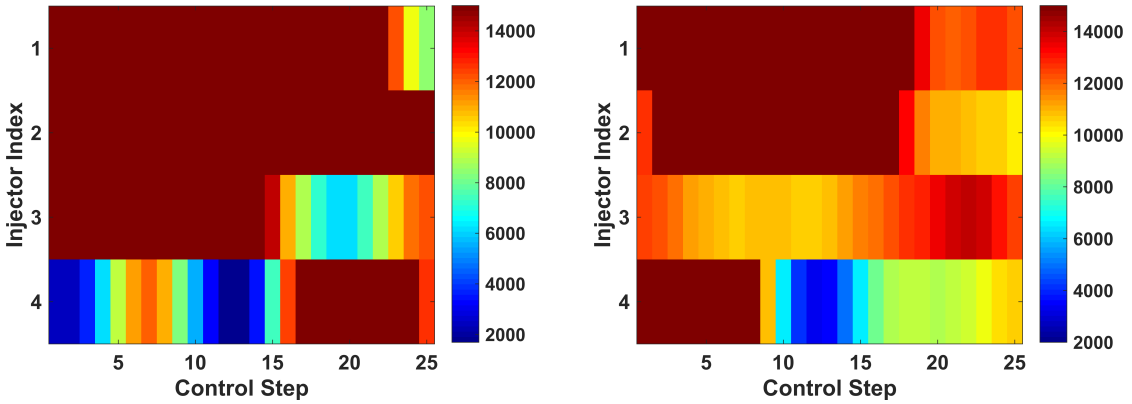


Figure 5.1: NPV versus forward runs, red stars represent CMG results; blue pluses represent INSIM-FT-3D results.

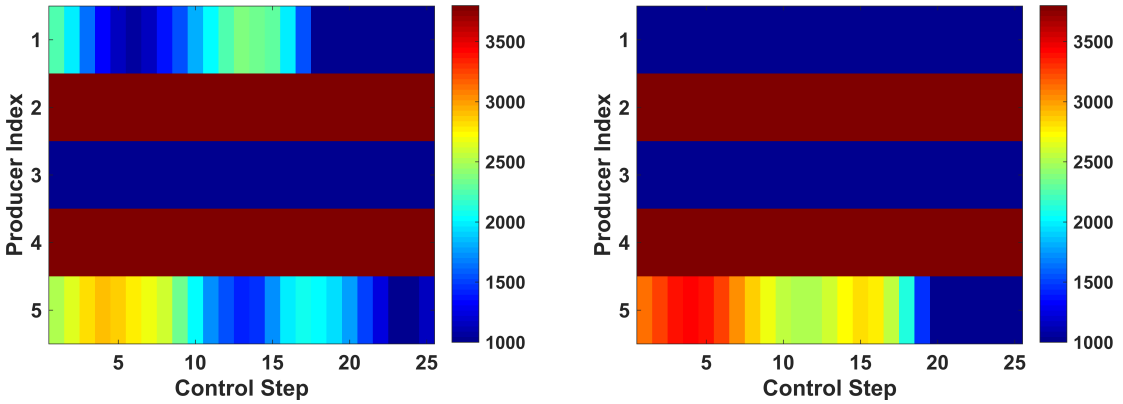


(a) INSIM-FT-3D (b) CMG

Figure 5.2: Comparison of optimal injection rates obtained with INSIM-FT-3D and CMG.

Table 5.1: Comparison of NPV obtained with different scenarios. INSIM-FT-3D/CMG refers to the NPV values obtained by inputting the INSIM-FT-3D generated initial and optimal controls into the CMG true model; Simplex Opt refers to the results obtained by optimizing well controls directly with CMG.

	Initial, 10^6 USD	Optimal, 10^6 USD
INSIM-FT-3D	2,060	2,390
INSIM-FT-3D/CMG	2,172	**2,406**
Simplex Opt	2,172	**2,495**



(a) INSIM-FT-3D (b) CMG

Figure 5.3: Comparison of optimal production BHP obtained with INSIM-FT-3D and CMG.

used one realization out of 104 as the true model to generate the production history for history matching with INSIM-FT-3D, with a history matching period equal to 3,600 days. Here, the same example is presented to test the performance of INSIM-FT-3D for production optimization, where optimization is done based on the history-matched INSIM-FT-3D model that gives the best data-match, m_{best}. The production optimization period lasts for 3,600 days, which is subsequent to the first 3,600 days of the history matching period. The number of control steps is 30 with each control step equal to 120 days in length. Considering that the total number of wells is 30, namely, 10 injectors and 20 producers, the total number of control variables is $30 \times 30 = 900$. The control variables for injectors are the water injection rates on control steps and for producers, the optimization variables are the BHPs on control steps with the well indices of producers computed with Eq. 5.13 based on m_{best}. The upper and lower bounds, respectively, for all BHP controls are 2,465 psi and 14.7 psi, respectively, where 14.7 psi is the lowest BHP used in producers during the history matching period and 2,465 psi is the initial reservoir pressure. The controls of injection rates are constrained to lie in the interval of [0, 8,000] RB/day. The initial guess of optimal controls are given by the controls at the end of history matching period, with injection rates of 4,000 RB/day for all the injectors. The initial guess of producer BHP varies from producer to producer; see Table 5.2.

Table 5.2: Initial guesses of optimal controls for producer BHP.

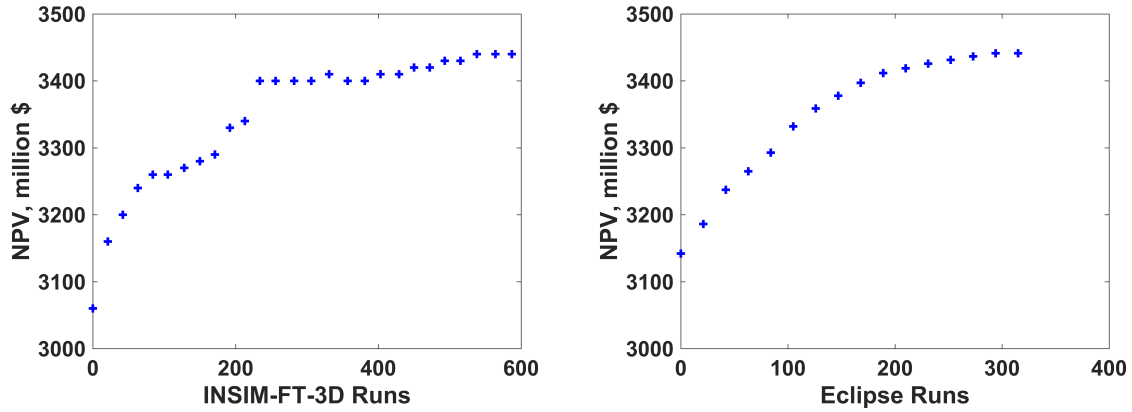
Well Name	Initial BHP, psi
P1	720.6593
P2	1694.343
P3	870.0442
P4	789.0767
P5	1061.085
P6	1289.202
P7	910.0908
P8	853.7908
P9	14.7
P10	14.7
P11	1692.818
P12	2465
P13	1689.314
P14	777.9531
P15	2465
P16	945.6213
P17	1198.156
P18	1490.184
P19	1252.633
P20	1646.328

Production optimization is performed with the INSIM-FT-3D history-matched model, m_{best} and the Eclipse true model and the comparison of NPV versus forward runs is shown in Fig. 5.4. It can be seen that, optimization with the two methods both yield an optimal NPV value around 3,450 million USD. To further validate the results obtained with INSIM-FT-3D, the optimal well controls is applied in the Eclipse true reservoir simulation model and the results are shown in Table 5.3. The results show that the optimal NPV obtained by optimization directly with Eclipse is slightly lower than is obtained by using the optimal well controls obtained with INSIM-FT-3D in the true model, which means that using simplex gradient with Eclipse as the forward model resulted in an estimated optimal NPV which is at least slightly suboptimal. In terms of computational efficiency, INSIM-FT-3D required 44 minutes for the entire optimization procedure whereas optimization with Eclipse requires 10.8 hours in total.

For the purpose of comparison, the optimal well controls generated with the true

Eclipse model and the INSIM-FT-3D model are shown in Figs. 5.5 and 5.6. The optimal well controls obtained with the two methods are qualitatively similar for some wells, such as P1, P9, P14 and P20, whereas are quite different for some wells. There are two possible reasons for such differences in well controls: (1) the optimization problem has multiple solutions for this field-scale problem, which all can give similar value of NPV (van Essen et al., 2011; Chen et al., 2012); (2) INSIM-FT-3D is a simplified proxy which cannot exactly reproduce the entire physics of the full-scale model.

Fig. 5.7 shows the optimized oil saturation fields in the first layer at the end of the reservoir life obtained with Eclipse (Fig. 5.7a) compared to those obtained with INSIM-FT-3D (Fig. 5.7b). The results show that, even though the optimal well controls with two methods are quite different for some wells, the final oil saturation fields are almost identical.

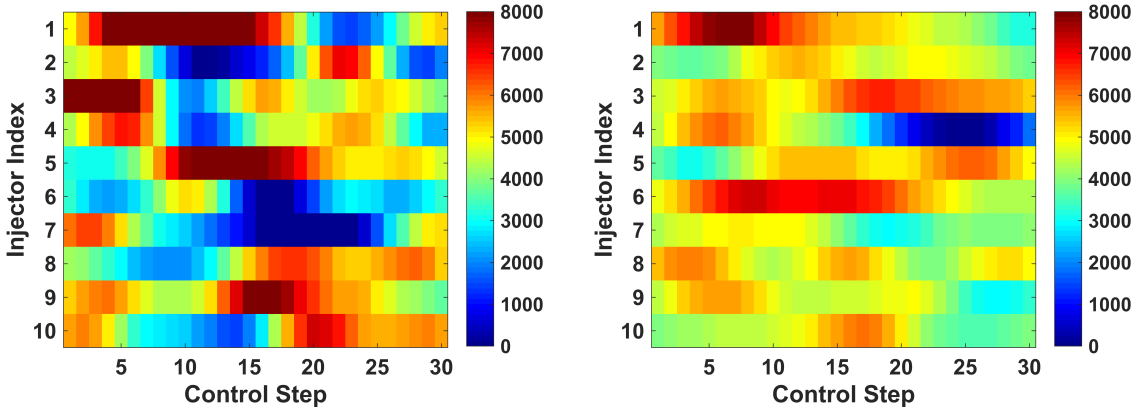


(a) INSIM-FT-3D (b) Eclipse

Figure 5.4: Comparison of optimal NPV values versus forward runs obtained with INSIM-FT-3D and Eclipse, Brugge example.

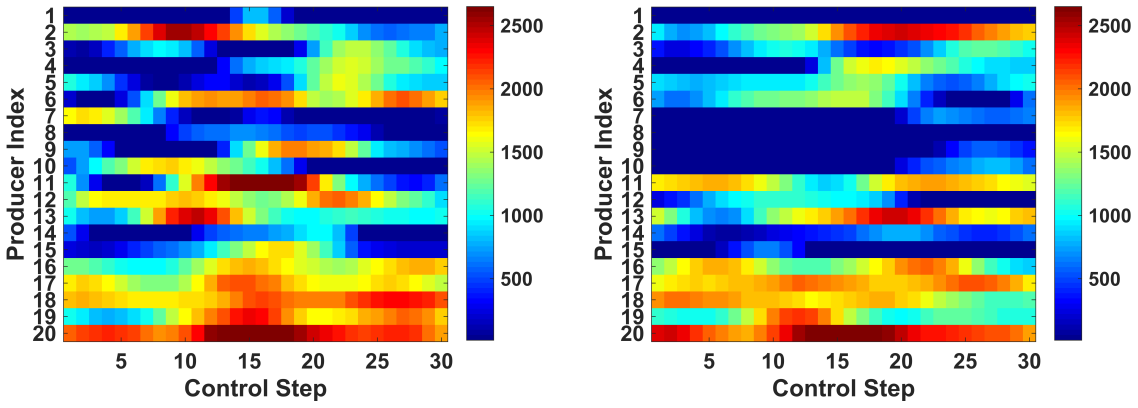
Table 5.3: Comparison of NPV obtained with different scenarios. INSIM-FT-3D/Eclipse refers to the NPV values obtained by inputting the INSIM-FT-3D generated initial and optimal controls into the Eclipse true model; Simplex Opt refers to the results obtained by optimizing well controls directly with Eclipse.

	Initial, 10^6 USD	Optimal, 10^6 USD
INSIM-FT-3D	3,058	3,475
INSIM-FT-3D/Eclipse	3,142	**3,532**
Simplex Opt	3,142	**3,444**



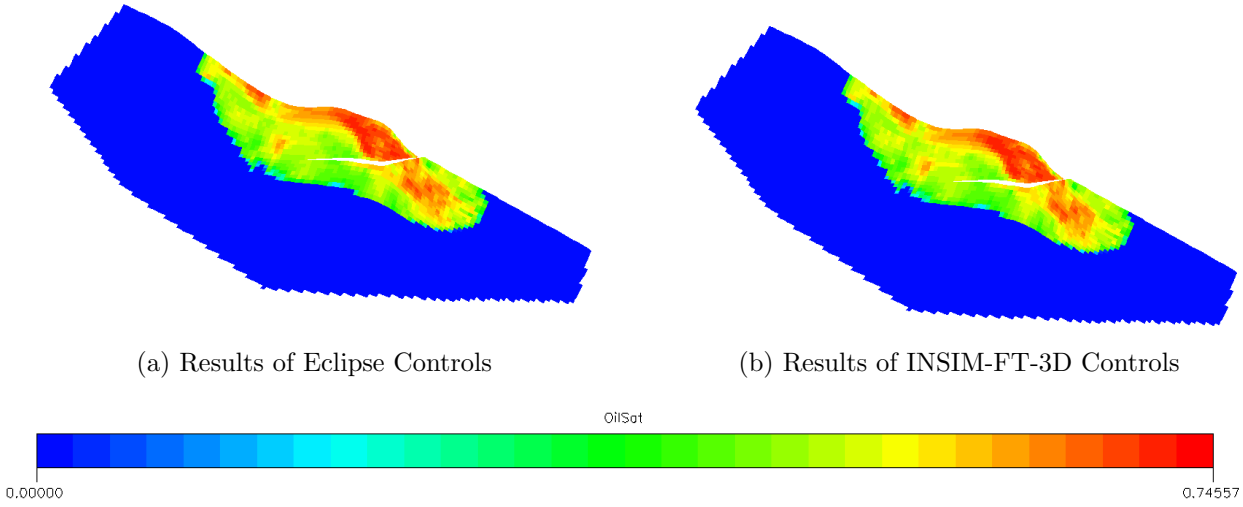
(a) INSIM-FT-3D (b) Eclipse

Figure 5.5: Comparison of optimal injection rates obtained with INSIM-FT-3D and Eclipse, Brugge example.



(a) INSIM-FT-3D (b) Eclipse

Figure 5.6: Comparison of optimal production BHP obtained with INSIM-FT-3D and Eclipse, Brugge example.



(a) Results of Eclipse Controls

(b) Results of INSIM-FT-3D Controls

Figure 5.7: Optimized oil saturation field.

CHAPTER 6

DISCUSSION AND CONCLUSIONS

In this work, we first develop a new data-driven model, the interwell numerical simulation model with front tracking (INSIM-FT) to history match, predict and optimize a single-layered reservoir under waterflooding. The new model overcomes the two primary defects of the original data-driven model, INSIM, which assumes relative permeabilities are known a priori and also used an ad hoc method for saturation computation. Then we extend INSIM-FT to allow for three-dimensional flow with gravity effects. The extended 3D model is referred to as INSIM-FT-3D. We show that history matching and production optimization can be successfully performed with INSIM-FT-3D.

INSIM-FT removes the deficiencies of INSIM by replacing the saturation calculation in INSIM by a front-tracking procedure based on the semi-analytical solution of Riemann and Cauchy problems and incorporating parameters of power-law relative permeabilities as additional model parameters. Also, unlike INSIM, INSIM-FT adds additional flow paths (stream tubes) via imaginary wells (additional nodes) to increase fidelity. Adding imaginary wells, however, requires the development of the new method to determine the interwell connectivity which is presented in this research. In addition, we show that when the interior-point optimizer used for history matching in the original INSIM algorithm is replaced by ES-MDA to obtain an modified procedure referred to as INSIM (ES-MDA), the history matching errors decrease but are still significantly larger, in most cases, than the errors obtained from INSIM-FT.

We apply INSIM-FT for two synthetic examples where we use a reservoir simulation model to represent the true reservoir in order to generate production data for history-matching as well as predictions of future reservoir performance. The computational results

indicate that INSIM-FT gives better history matches, lower prediction errors and better estimates of reservoir connectivity than are obtainable with the original INSIM where history matching is done with a Matlab interior point optimizer or with INSIM (ES-MDA) which uses ES-MDA for history matching. This is the expected result based on the fact that INSIM-FT incorporates more of the correct physics of flow than does INSIM. The results demonstrate that INSIM-FT can detect the presence of a fault whereas INSIM cannot.

We also implement a modern capacitance resistance model (CRM) and compare its performance with INSIM-FT. Overall, INSIM-FT outperforms CRM in terms of the quality of the history match and determining reservoir connectivity. For the first example presented in Chapter 2, however, the prediction error of CRM is almost zero and is much lower than the history-matching error. This CRM result appears to be an anomaly as we expect the history-matching error to be smaller than the prediction error which is always the case for the INSIM, INSIM (ES-MDA) and INSIM-FT results obtained and is also the case for the CRM results obtained for the second example in Chapter 2.

For production optimization with INSIM-FT, we aim to use the INSIM-FT model obtained by history-matching as a forward model to estimate the well controls that optimize the NPV of production for the remaining life of the reservoir. A method for estimating well indices from the history-matched INSIM-FT model is presented in order to enable using flowing wellbore pressures as well controls provided that data for the flowing wellbore pressures are available during the history-matching period. If wellbore pressure data are not available, only rate controls can be used when optimizing the NPV of production. We consider the same two synthetic examples for production optimization, which are tested for history matching with INSIM-FT. It is shown that when the optimal well controls generated using INSIM-FT as the forward predictive model during production optimization are input into the true reservoir model (reservoir simulation model), the NPV obtained is in good agreement with the estimate of NPV obtained from optimization based on the INSIM-FT model. Even for the very small synthetic examples considered, optimization using INSIM-FT as the forward model requires more than one order of magnitude less computational

time than is required when production optimization is done using the reservoir simulation model as the forward model. The production optimization methodology based on INSIM-FT was applied successfully to the same field example that is tested for history matching with INSIM-FT.

We also extend the INSIM-FT model from two dimensions to three dimensions. The new data-driven model, which is referred to as INSIM-FT-3D, includes gravitational effects. When gravity is included, the water fractional flow curve may be no longer S-shaped and generally may have two inflection points. In order to solve the Riemann problem when the fractional flow function has more than one inflection point, we introduced a convex-hull method. We develop an automated workflow by using the Mitchell's best-candidate algorithm and Delaunay triangulation with a modification in order to generate the imaginary wells and the connection map for INSIM-FT-3D. By using this workflow, we avoid the manual construction of the connection map for a large-scale problem with many wells of arbitrary trajectories. As demonstrated by a synthetic example, INSIM-FT-3D can handle wells with different orientations and inclinations with multiple perforated well segments. For the field example considered, the results suggest that including gravity improves the history match and future reservoir performance predictions. The Brugge example demonstrates the applicability of INSIM-FT-3D for a field-scale problem with many wells.

We design a workflow to optimize well controls for a waterflooding reservoir based on using the history-matched INSIM-FT-3D model as the forward model. Two three-dimensional test cases including a Brugge field-scale reservoir are tested. From the computational results, we show that the NPV calculated with the true reservoir model using the optimal well controls obtained with INSIM-FT-3D is quite close to the INSIM-FT-3D-generated optimal NPV. By inputting the optimal well controls generated with INSIM-FT-3D in the true model, the obtained NPV is comparable with the optimal NPV generated by optimization directly with the true model. However, the computational cost required to optimize with INSIM-FT-3D is an order of magnitude less than what is required to optimize NPV using the true reservoir model and a commercial reservoir simulator as the forward

model. For the two examples presented, INSIM-FT-3D only takes less than one tenth of the computational time that is required for optimization with a full-scale reservoir simulation model.

Based on this work, the following conclusions are warranted.

1. INSIM-FT is more robust and flexible for assisted history matching than the original INSIM, by providing a more rigorous way for saturation computation, introducing imaginary wells and adding the option for matching relative permeabilities.

2. INSIM-FT gives a more accurate estimates of injection well allocation than are generated with CRM and INSIM.

3. INSIM-FT generally provides better history matches and predictions than does CRM.

4. INSIM-FT-3D is viable for 3D reservoirs with gravity effects.

5. In INSIM-FT-3D, the Buckley-Leverett problem with a fractional flow curve having more than one inflection point can be successfully solved using a convex-hull method.

6. The INSIM-FT-3D model gives a slightly better history match and predictions for a real field problem than are obtained with the original INSIM-FT model.

7. The optimal well controls obtained from INSIM-FT or INSIM-FT-3D when inputting into the true model yield an estimated optimal NPV within 5% of the NPV generated by directly optimizing the well controls using the true reservoir simulation.

8. The forward run of an INSIM-FT or INSIM-FT-3D model requires the computational cost that is at least an order of magnitude less than what is required by running a full-scale reservoir simulation model.

6.1 Future Work

The future work to improve the INSIM-FT model is to add the capability to history match BHP data, which is not yet implemented in our INSIM-FT or INSIM-FT-3D software.

Some studies have been done to improve the accuracy and resolution of the INSIM-FT model; however, the history matching quality in terms of matching the BHP data is not significantly improved. Even though we have not found a good solution to solve the BHP-matching problem, it is worthwhile to show the studies we have done so far, i.e., to compute the transmissibility along each connection in a more rigorous way. Currently, the transmissibility is computed using Eq. A-19, which is repeated in the following equation:

$$T_{i,j}^{n-1} = 1.127 \times 10^{-3} k_{i,j} A_{i,j} \frac{\lambda_t(S_{w,i,j}^{n-1})}{L_{i,j}} = T_{i,j}^{0} \frac{\lambda_t(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^{0})}, \tag{6.1}$$

where $\lambda_t(S_{w,i,j})$ is determined by upstream weighting, i.e., if node i is upstream of node j, then $S_{w,i,j} = S_{w,i}$; otherwise, $S_{w,i,j} = S_{w,j}$. By neglecting gravity effects, the total flow rate along connection (i, j) is given by

$$q_{t,i,j}^{n} = T_{i,j}^{n-1}(p_j^n - p_i^n) = 1.127 \times 10^{-3} k_{i,j} A_{i,j} \frac{\lambda_t(S_{w,i,j}^{n-1})}{L_{i,j}(p_j^n - p_i^n)}. \tag{6.2}$$

However, from the Darcy law, the total flow rate along connection (i, j) is given by

$$q_{t,i,j} = -1.127 \times 10^{-3} k_{i,j} A_{i,j} \lambda_t(S_w(x)) \frac{dp}{dx}. \tag{6.3}$$

Following the INSIM-FT-3D formulation as Eq. 2.1, we consider the water saturation in Eq. 6.3 is provided at time level $n-1$ and pressure is at time level n, then Eq. 6.3 is rewritten as

$$q_{t,i,j}^{n} = -1.127 \times 10^{-3} k_{i,j} A_{i,j} \lambda_t(S_w^{n-1}(x)) \frac{dp^n}{dx}. \tag{6.4}$$

By comparing Eq. 6.2 with Eq. 6.3, we can see that Eq. 6.2 gives a good estimate of $q_{t,i,j}^{n}$ only if $\lambda_t(S_{w,i,j}^{n-1}) \frac{p_j^n - p_i^n}{L_{i,j}}$ is a good estimate of $\lambda_t(S_w^{n-1}(x)) \frac{dp^n}{dx}$, which we believe is a very strong assumption. Next, we will derive a more rigorous method to compute $T_{i,j}$ rather than using Eq. 6.1. Rearranging Eq. 6.3 yields

$$-\int_{p_j}^{p_i} dp = p_j - p_i = \frac{q_{t,i,j}}{1.127 \times 10^{-3} k_{i,j} A_{i,j}} \int_{x_j}^{x_i} \lambda_t^{-1}(S_w(x))dx, \tag{6.5}$$

Where x_i and x_j, respectively, represent the location of node i and j. As Eq. 2.1, in Eq. 6.5, we consider the pressures of node i and j are at level n and the saturation profile of connection (i, j) is provided at time level $n-1$, then Eq. 6.5 is rewritten as

$$p_j^n - p_i^n = \frac{q_{t,i,j}^n}{1.127 \times 10^{-3} k_{i,j} A_{i,j}} \int_{x_j}^{x_i} \lambda_t^{-1}(S_w^{n-1}(x))dx. \tag{6.6}$$

Since we use front tracking to solve the saturation profile along connection (i, j), $S_w^{n-1}(x)$ for $0 < x < L_{i,j}$ is obtained as a piecewise constant function. By assuming that the piecewise constant function has n_s constant states which are represented by

$$[S_{w,1}, S_{w,2}, \ldots S_{w,n_s}],$$

and each state has a length of Δx_k, $k = 1, 2, \ldots n_s$, $\int_{x_j}^{x_i} \lambda_t^{-1}(S_w(x)^{n-1})dx$ in Eq. 6.6 can be calculated as

$$\int_{x_j}^{x_i} \lambda_t^{-1}(S_w^{n-1}(x))dx = \sum_{k=1}^{n_s} \frac{\Delta x_k}{\lambda_t(S_{w,k})}. \tag{6.7}$$

Substituting Eq. 6.7 into Eq. 6.6 yields

$$p_j^n - p_i^n = \frac{q_{t,i,j}^n}{1.127 \times 10^{-3} k_{i,j} A_{i,j}} \sum_{k=1}^{n_s} \frac{\Delta x_k}{\lambda_t(S_{w,k})}. \tag{6.8}$$

By defining

$$\overline{\lambda}_t^{n-1} = \frac{L_{i,j}}{\sum_{k=1}^{n_s} \frac{\Delta x_k}{\lambda_t(S_{w,k})}},$$

Eq. 6.8 is simplified as

$$q_{t,i,j}^n = \frac{1.127 \times 10^{-3} k_{i,j} A_{i,j}}{L_{i,j}} \overline{\lambda}_t^{n-1} (p_j^n - p_i^n). \tag{6.9}$$

By comparing Eq. 6.9 with Eq. 6.2, $T_{i,j}^{n-1}$ can be computed as

$$T_{i,j}^{n-1} = \frac{1.127 \times 10^{-3} k_{i,j} A_{i,j}}{L_{i,j}} \overline{\lambda}_t^{n-1} = T_{i,j}^{0} \frac{\overline{\lambda}_t^{n-1}}{\lambda_t(S_{w,i,j}^{0})}, \tag{6.10}$$

From the above derivations, it is shown that using Eq. 6.10 to compute $T_{i,j}^{n-1}$ has more theoretical basis than using Eq. 6.1 under the situation that gravity effects can be neglected. In our INSIM-FT-3D software, we have added the option to compute $T_{i,j}^{n-1}$ using Eq. 6.10, however, the history matching quality using this option for pressure data match is not significantly improved as compared to that obtained using the original formulation. Part of the reason might be due to the lack of consideration of gravity effects in the above derivations, or there is still some other key physics we have not captured in our INSIM-FT-(3D) models to enable the pressure match.

BIBLIOGRAPHY

Albertoni, A. and L. W. Lake, Inferring connectivity only from well-rate fluctuations in water floods, *SPE Reservoir Evaluation & Engineering*, **6**(1), 6–16, 2003.

Aziz, K. and A. Settari, *Petroleum reservoir simulation*, Chapman & Hall, 1979.

Brouwer, D. and J. Jansen, Dynamic optimization of water flooding with smart wells using optimal control theory, *SPE Journal*, **9**, 391–402, 2004.

Cao, F., *Development of a two-phase flow coupled capacitance resistance model*, Ph.D. thesis, The University of Texas at Austin, Austin, Texas, 2014.

Cao, F., H. Luo, and L. W. Lake, Development of a fully coupled two-phase flow based capacitance resistance model (CRM),, in *Proceedings of SPE Improved Oil Recovery Symposium, Tulsa, Oklahoma, USA, 12-16 April*, SPE-169485-MS, 2014.

Cao, F., H. Luo, and L. W. Lake, Oil-rate forecast by inferring fractional-flow models from field data with koval method combined with the capacitance/resistance model, *SPE Reservoir Evaluation & Engineering*, **18**(04), 534–553, 2015.

Cardoso, M. A. and L. J. Durlofsky, Use of reduced-order modeling procedures for production optimization, *SPE Journal*, **15**(02), 426–435, 2010.

Castellini, A., H. Gross, Y. Zhou, J. He, and W. Chen, An iterative scheme to construct robust proxy models, in *Proceedings of ECMOR XII-12th European Conference on the Mathematics of Oil Recovery*, 2010.

Chen, B., *A Stochastic Simplex Approximate Gradient for Production Optimization of WAG and Continuous Water Flooding*, Ph.D. thesis, The University of Tulsa, Tulsa, Oklahoma, 2017.

Chen, B., R.-M. Fonseca, O. Leeuwenburgh, and A. C. Reynolds, Minimizing the risk in the robust life-cycle production optimization using stochastic simplex approximate gradient, *Journal of Petroleum Science and Engineering*, **153**, 331–344, 2017.

Chen, B. and A. C. Reynolds, Ensemble-based optimization of the water-alternating-gas-injection process, *SPE Journal*, **21**(03), 786–798, 2016.

Chen, B. and A. C. Reynolds, CO2 water-alternating-gas injection for enhanced oil recovery: Optimal well controls and half-cycle lengths, *Computers & Chemical Engineering*, **113**, 44 – 56, 2018.

Chen, C., *Adjoint-gradient-based production optimization with the augmented Lagrangian method*, Ph.D. thesis, The University of Tulsa, Tulsa, Oklahoma, 2011.

Chen, C., G. Li, and A. C. Reynolds, Robust constrained optimization of short and long-term NPV for closed-loop reservoir management, *SPE Journal*, **17**(3), 849–864, 2012.

Chen, Y. and D. Oliver, Ensemble-based closed-loop optimization on brugge field, SPE118926, 2009.

Chen, Y. and D. S. Oliver, Ensemble randomized maximum likelihood method as an iterative ensemble smoother, *Mathematical Geosciences*, **44**(1), 1–26, 2012.

Chen, Y., D. S. Oliver, and D. Zhang, Efficient ensemble-based closed-loop production optimization, *SPE Journal*, **14**(4), 634–645, 2009.

De Berg, M., M. Van Kreveld, M. Overmars, and O. C. Schwarzkopf, Computational geometry, in *Computational Geometry*, pp. 1–17, Springer, 2000.

de Montleau, P., A. Cominelli, D. R. K. Neylon, I. Pallister, O. Tesaker, and I. Nygard, Production optimization under constraints using adjoint gradients, in *Proceedings of the 10th European Conference on the Mathematical Oil Recovery - Amsterdam, 4-7 September*, SPE 109805, 2006.

Devadoss, S. L. and J. O'Rourke, *Discrete and computational geometry*, Princeton University Press, 2011.

Do, S. T. and A. C. Reynolds, Theoretical connections between optimization algorithms based on an approximate gradient, *Computational Geosciences*, **17**(6), 959–973, 2013.

Dunbar, D. and G. Humphreys, A spatial data structure for fast Poisson-disk sample generation, *ACM Transactions on Graphics (TOG)*, **25**(3), 503–508, 2006.

Edelsbrunner, H., T. S. Tan, and R. Waupotitsch, An O(n^2log(n)) time algorithm for the minmax angle triangulation, *SIAM Journal on Scientific and Statistical Computing*, **13**(4), 994–1008, 1992.

Emerick, A. A. and A. C. Reynolds, History matching time-lapse seismic data using the ensemble Kalman filter with multiple data assimilations, *Computational Geosciences*, **16**(3), 639–659, 2012.

Emerick, A. A. and A. C. Reynolds, Ensemble smoother with multiple data assimilations, *Computers & Geosciences*, **55**, 3–15, 2013a.

Emerick, A. A. and A. C. Reynolds, History-matching production and seismic data in a real field case using the ensemble smoother with multiple data assimilation, in *Proceedings of the SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, 18-20 February*, SPE-163645-MS, 2013b.

Evensen, G., Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *Journal of Geophysical Research*, **99**(C5), 10,143–10,162, 1994.

Fonseca, R., O. Leeuwenburgh, P. V. den Hof, and J. D. Jansen, Improving the ensemble optimization procedure through covariance matrix adaptation (CMA-EnOpt), in *Proceedings of the SPE Reservoir Simulation Symposium*, SPE 163657, 2013.

Fonseca, R., O. Leeuwenburgh, E. D. Rossa, P. V. den Hof, and J. D. Jansen, Ensemble-based multi-objective optimization of on-off control devices under geological uncertainty, in *Proceedings of SPE Reservoir Simulation Symposium*, SPE 173268, Montgomery, Texas, USA, 20-22 February, 2015.

Fonseca, R. M., B. Chen, J. D. Jansen, and A. C. Reynolds, A stochastic simplex approximate gradient (StoSAG) for optimization under uncertainty, *International Journal for Numerical Methods in Engineering*, **Online First**, 2016.

Gao, G., G. Li, and A. C. Reynolds, A stochastic algorithm for automatic history matching, in *Proceedings of the SPE Annual Technical Conference and Exhibition*, SPE 90065, 2004.

Gao, G. and A. C. Reynolds, An improved implementation of the LBFGS algorithm for automatic history matching, *SPE Journal*, **11**(1), 5–17, 2006.

Gao, G., J. C. Vink, C. Chen, F. O. Alpak, and K. Du, A parallelized and hybrid data-integration algorithm for history matching of geologically complex reservoirs, *SPE Journal*, **21**(06), 2–155, 2016.

Gentil, P., *The Use of Multilinear Regression Models in Patterned Waterfloods: Physical Meaning of the Regression Coefficients*, Master's thesis, University of Texas at Austin, Austin, Texas, 2005.

Gildin, E., M. Ghasemi, A. Romanovskay, and Y. Efendiev, Nonlinear complexity reduction for fast simulation of flow in heterogeneous porous media, in *Proceedings of SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, 18-20 February*, SPE-163618-MS, 2013.

Gu, Y. and D. S. Oliver, An iterative ensemble Kalman filter for multiphase fluid flow data assimilation, *SPE Journal*, **12**(4), 438–446, 2007.

Guo, Z., C. Chen, G. Gao, R. Cao, R. Li, and C. Liu, EUR assessment of unconventional assets using machine learning and distributed computing techniques, in *Proceedings*

of SPE/AAPG/SEG Unconventional Resources Technology Conference, Austin, Texas, USA, 24-26 July, URTEC-2659996-MS, Unconventional Resources Technology Conference (URTEC), 2017a.

Guo, Z., C. Chen, G. Gao, and J. Vink, Applying support vector regression to reduce the effect of numerical noise and enhance the performance of history matching, in *Proceedings of SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA, 9-11 October*, SPE-187430-MS, Society of Petroleum Engineers, 2017b.

Guo, Z., A. C. Reynolds, and H. Zhao, A physics-based data-driven model for history-matching, prediction and characterization of waterflooding performance, in *Proceedings of the SPE Reservoir Simulation Conference*, SPE-182660-MS, 2017c.

Guo, Z., A. C. Reynolds, and H. Zhao, A physics-based data-driven model for history matching, prediction, and characterization of waterflooding performance, *SPE Journal*, in press 2018a.

Guo, Z., A. C. Reynolds, and H. Zhao, Waterflooding optimization with the INSIM-FT data-driven model, *Computational Geosciences*, in press 2018b.

He, J. and L. J. Durlofsky, Reduced-order modeling for compositional simulation by use of trajectory piecewise linearization, *SPE Journal*, **19**(05), 858–872, 2014.

He, J., J. Xie, X.-H. Wen, and W. Chen, Improved proxy for history matching using proxy-for-data approach and reduced order modeling, in *Proceedings of SPE Western Regional Meeting, Garden Grove, California, USA, 27-30 April*, SPE-174055-MS, 2015.

Heffer, K. J., R. J. Fox, C. A. McGill, and N. Koutsabeloulis, Novel techniques show links between reservoir flow directionality, earth stress, fault structure and geomechanical changes in mature waterfloods, *SPE Journal*, 1997.

Holden, H., L. Holden, and R. Høegh-Krohn, A numerical method for first order nonlinear

scalar conservation laws in one-dimension, *Computers & Mathematics with Applications*, **15**(6), 595–602, 1988.

Iglesias, M. A. and C. Dawson, The regularizing Levenberg–Marquardt scheme for history matching of petroleum reservoirs, *Computational geosciences*, **17**(6), 1033–1053, 2013.

IMEX, *User Guide*, Computer Modelling Group Ltd., 2010.

Jansen, F. E. and M. G. Kelkar, Non-stationary estimation of reservoir properties using production data, in *SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA, 5-8 October*, SPE-38729-MS, 1997.

Jansen, J., D. Brouwer, G. Naevdal, and C. van Kruijsdijk, Closed-loop reservoir management, *First Break*, **23**, 43–48, 2005.

Jansen, J. D., S. D. Douma, D. R. Brouwer, P. M. J. V. den Hof, and A. W. Heemink, Closed-loop reservoir management, in *Proceedings of the SPE Reservoir Simulation Symposium, The Woodlands, Texas, 2–4 February*, SPE 119098, 2009.

Juanes, R. and T. W. Patzek, Multiscale numerical modeling of three-phase flow, in *Proceedings of the SPE Annual Technical Conference and Exhibition*, SPE-84369-MS, 2003.

Kahrobaei, S., G. Van Essen, J. Van Doren, P. Van den Hof, J. Jansen, et al., Adjoint-based history matching of structural models using production and time-lapse seismic data, in *SPE Reservoir Simulation Symposium*, Society of Petroleum Engineers, 2013.

Kraaijevanger, J. F. B. M., P. J. P. Egberts, J. R. Valstar, and H. W. Buurman, Optimal waterflood design using the adjoint method, in *Proceedings of the SPE Reservoir Simulation Symposium*, SPE-105764-MS, 2007.

Lake, L. W., X. Liang, T. F. Edgar, A. Al-Yousef, M. Sayarpour, and D. B. Weber, Optimization of oil production based on a capacitance model of production and injection rates, in *Proceedings of hydrocarbon economics and evaluation symposium, Dallas, Texas, USA, 1-3 April*, SPE-107713-MS, 2007.

Landa, J. L. and B. Güyagüler, A methodology for history matching and the assessment of uncertainties associated with flow prediction, in *Proceedings of SPE Annual Technical Conference and Exhibition, Denver, Colorado, USA, 5-8 October*, SPE-84465-MS, 2003.

Le, D. H., A. A. Emerick, and A. C. Reynolds, An adaptive ensemble smoother with multiple data assimilation for assisted history matching, *SPE Journal*, **21**(06), 2,195–2,207, 2016.

Lerlertpakdee, P., B. Jafarpour, and E. Gildin., Efficient production optimization with flow-network models., *SPE Journal*, **19**, 1083–1095, 2014.

Li, G. and A. C. Reynolds, Iterative ensemble Kalman filters for data assimilation, *SPE Journal*, **14**(3), 496–505, 2009.

Li, R., A. C. Reynolds, and D. S. Oliver, History matching of three-phase flow production data, *SPE Journal*, **8**(4), 328–340, 2003.

Lie, K. A. and R. Juanes, A front-tracking method for the simulation of three-phase flow in porous media, *Computational Geosciences*, **9**(1), 29–59, 2005.

Liu, D. and J. Nocedal, On the limited memory BFGS method for large scale optimization, *Mathematical Programming*, **45**, 503–528, 1989.

Liu, X. and A. C. Reynolds, Augmented Lagrangian method for maximizing expectation and minimizing risk for optimal well control problems with nonlinear constraints, *SPE Journal*, **21**(5), 1830–1842, 2016a.

Liu, X. and A. C. Reynolds, Gradient-based multi-objective optimization with applications to waterflooding optimization, *Computational Geosciences*, **20**(3), 677–693, 2016b.

Lorentzen, R. J., A. M. Berg, G. Nævdal, and E. H. Vefring, A new approach for dynamic optimization of waterflooding problems, in *Proceedings of the SPE Intelligent Energy Conference and Exhibition*, SPE 99690, 2006.

Markovinovic, R., E. Geurtsen, T. Heijn, and J. Jansen, Generation of low-order reservoir models using pod, emperical grammians and subspace identification, in *Proceedings of ECMOR VIII-8th European Conference on the Mathematics of Oil Recovery*, 2002.

MathWorks, The language of technical computing., *MATLAB*, 2004.

Mitchell, D. P., Spectrally optimal sampling for distribution ray tracing, in *ACM SIGGRAPH Computer Graphics*, vol. 25, pp. 157–164, ACM, 1991.

Nævdal, G., T. Mannseth, and E. H. Vefring, Near-well reservoir monitoring through ensemble Kalman filter, in *Proceedings of the SPE/DOE Improved Oil Recovery Symposium, 13–17 April*, SPE 75235, 2002.

Nash, S. G. and A. Sofer, *Linear and nonlinear programming*, Ohio: McGraw-Hill Science/Engineering/Math, 1996.

Nguyen, A. P., *Capacitance Resistance Modeling for Primary Recovery, Waterflood and Water-CO2 Flood*, Ph.D. thesis, University of Texas at Austin, Austin, Texas, 2012.

Oleinik, O. A., Discontinuous solutions of non-linear differential equations, *Uspekhi Matematicheskikh Nauk*, **12**(3), 3–73, 1957.

Oliveira, D. F. and A. C. Reynolds, An adaptive hierarchical multiscale algorithm for estimation of optimal well controls, *SPE Journal*, **19**(05), 909–930, 2014.

Oliver, D. S., N. He, and A. C. Reynolds, Conditioning permeability fields to pressure data, in *Proceedings of the European Conference for the Mathematics of Oil Recovery*, 1996.

Peaceman, D. W., Interpretation of well-block pressures in numerical reservoir simulation (includes associated paper 6988), *Society of Petroleum Engineers Journal*, **18**(03), 183–194, 1978.

Peaceman, D. W., Interpretation of well-block pressures in numerical reservoir simulation with nonsquare grid blocks and anisotropic permeability, *Society of Petroleum Engineers Journal*, **23**(03), 531–543, 1983.

Peters, L., R. Arts, G. Brouwer, C. Geel, S. Cullick, R. Lorentzen, Y. Chen, K. Dunlop, F. Vossepoel, R. Xu, P. Sarma, A. Alhuthali, and A. Reynolds, Results of the Brugge benchmark study for flooding optimisation and history matching, *SPE Reservoir Evaluation & Engineering*, **13**(3), 391–405, 2010.

Powell, M. J., Least Frobenius norm updating of quadratic models that satisfy interpolation conditions, *Math programming*, **100**, 183–215, 2004.

Refunjol, B. T., *Reservoir characterization of North Buck Draw field based on tracer response and production/injection analysis*, Master's thesis, University of Texas at Austin, 1996.

Reynolds, A. C., N. He, and D. S. Oliver, Reducing uncertainty in geostatistical description with well testing pressure data, in *Reservoir Characterization–Recent Advances*, (edited by R. A. Schatzinger and J. F. Jordan), pp. 149–162, American Association of Petroleum Geologists, 1999.

Reynolds, A. C., R. Li, and D. S. Oliver, Simultaneous estimation of absolute and relative permeability by automatic history matching of three-phase flow production data, *J. Canadian Pet. Tech.*, **43**(3), 37–46, 2004.

Reynolds, A. C., M. Zafari, and G. Li, Iterative forms of the ensemble Kalman filter, in *Proceedings of 10th European Conference on the Mathematics of Oil Recovery, Amsterdam, 4–7 September*, 2006.

Sarma, P., K. Aziz, and L. Durlofsky, Implementation of adjoint solution for optimal control of smart wells, SPE-92864-ms, 2005.

Sayarpour, M., *Development and Application of Capacitance-Resistive Models to Water/C02 Flood*, Ph.D. thesis, University of Texas at Austin, Austin, Texas, 2008.

Schlumberger, Eclipse reference manual (2013.1), 2013a.

Schlumberger, Frontsim user guide (2013.1), 2013b.

Smola, A. and V. Vapnik, Support vector regression machines, *Advances in neural information processing systems*, **9**, 155–161, 1997.

Tarantola, A., *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, USA, 2005.

Thiele, M. R. and R. P. Batycky, Using streamline-derived injection efficiencies for improved waterflood management, *SPE Reservoir Evaluation & Engineering*, **9**(02), 187–196, 2006.

van Doren, J. F., R. Markovinović, and J. D. Jansen, Reduced-order optimal control of water flooding using proper orthogonal decomposition, *Computational Geosciences*, **10**(1), 137–158, 2006.

van Essen, G., P. V. den Hof, and J. Jansen, Hierarchical optimization of oil production from petroleum reservoirs, in *Workship on Data Assimilation and Reservoir Optimization, 20 January*, Technical University of Delft, 2009a.

van Essen, G., M. Zandvliet, P. V. den Hof, O. Bosgra, and J. Jansen, Robust waterflooding optimization of multiple geological scenarios, in *Proceedings of the SPE Annual Technical Conference and Exhibition*, SPE 84571-MS, 2006.

van Essen, G., M. Zandvliet, P. V. den Hof, O. Bosgra, and J. Jansen, Robust waterflooding optimization of multiple geological scenarios, SPE 102913, *SPE Journal*, **14**(1), 202–210, 2009b.

van Essen, G. M., P. M. J. V. den Hof, and J. D. Jansen, Hierarchical long-term and short-term production optimization, *SPE Journal*, **16**(1), 191–199, 2011.

van Leeuwen, P. J. and G. Evensen, Data assimilation and inverse methods in terms of a probabilistic formulation, *Monthly Weather Review*, **124**, 2898–2913, 1996.

Vogel, C. R., *Computational methods for inverse problems*, SIAM, 2002.

Weber, D. B., *The use of capacitance-resistance models to optimize injection allocation and well location in water floods*, Ph.D. thesis, The University of Texas at Austin, Austin, Texas, 2009.

Yan, X. and A. C. Reynolds, Optimization algorithms based on combining fd approximations and stochastic gradients compared with methods based only on a stochastic gradient, *SPE Journal*, **19**(05), 873–890, 2014.

Yousef, A. A., P. H. Gentil, J. L. Jensen, and L. W. Lake, A capacitance model to infer interwell connectivity from production and injection rate fluctuations, in *Proceedings of SPE Annual Technical Conference and Exhibition, Dallas, Texas, USA, 9-12 October*, SPE-95322-MS, 2005.

Yousef, A. A., P. H. Gentil, J. L. Jensen, and L. W. Lake, A capacitance model to infer interwell connectivity from production and injection rate fluctuations, *SPE Journal*, **9**(06), 630–646, 2006.

Yuan, B., R. G. Moghanloo, and D. Zheng, Analytical evaluation of nanoparticle application to mitigate fines migration in porous media, *SPE Journal*, 2016.

Zakirov, I. S., S. I. Aanonsen, E. S. Zakirov, and B. M. Palatnik, Optimizating reservoir performance by automatic allocation of well rates, in *Proceedings of the 5th European Conference on the Mathematical Oil Recovery - Leoben, Austria, 3-5 September*, 1996.

Zhao, H., C. Chen, S. Do, D. Oliveira, G. Li, and A. C. Reynolds, Maximization of a dynamic quadratic interpolation model for production optimization, *SPE Journal*, **18**, 1012–1025, 2013.

Zhao, H., Z. Kang, X. Zhang, H. Sun, L. Cao, and A. C. Reynolds, INSIM: A data-driven model for history matching and prediction for waterflooding monitoring and management with a field application, in *Proceedings of SPE Reservoir Simulation Symposium, Houston, Texas, USA, 23-25 February*, SPE-173213-MS, 2015.

Zhao, H., Z. Kang, X. Zhang, H. Sun, L. Cao, and A. C. Reynolds, A physics-based data-driven numerical model for reservoir history matching and prediction with a field application, *SPE Journal*, **21**(06), 2175–2194, 2016.

APPENDIX A

DERIVATION OF INSIM-FT

Considering the two-phase flow of oil and water and neglecting gravity and capillarity effects (Guo et al., 2018a), the material balance equation at the bulk volume controlled by well i, which is denoted by $V_{b,i}$, is given by

$$\frac{1}{5.615}\frac{\partial}{\partial t}\int_{V_{b,i}}(\rho_m\phi S_m)dv = 1.127\times10^{-3}\int_{S_i}(\rho_m\frac{k\cdot k_{rm}}{\mu_m}\frac{\partial p}{\partial n})dS + \rho_m q_{m,i}, \quad m=o,w, \quad \text{(A-1)}$$

where S_i in ft^2 denotes the surface area of bulk volume, $V_{b,i}$ in ft^3; ϕ is the porosity; S_m is the saturation; and μ_m is the viscosity in cp of m phase for $m=o,w$; k is the absolute permeability in md. Under the subsurface condition, ρ_m is the fluid density in lbm/RB and $q_{m,i}$ is the volumetric phase rate in RB/day injected/produced from the control volume i, for $m=o,w$, where a positive value of $q_{m,i}$ denotes injection and a negative value denotes production. Eliminating the integrand at the first term by taking the average property of this control volume i gives,

$$\begin{aligned}
&\frac{\partial}{\partial t}\int_{V_{b,i}}(\rho_m\phi S_m)dv \\
=&V_{b,i}\frac{\partial}{\partial t}(\rho_{m,i}\phi_i S_{m,i}) \\
=&\rho_{m,i}V_{b,i}\phi_i\left(\frac{\partial S_{m,i}}{\partial t}+S_{m,i}(c_r+c_m)\frac{\partial p_i}{\partial t}\right) \\
=&\rho_{m,i}V_{p,i}\left(\frac{\partial S_{m,i}}{\partial t}+S_{m,i}(c_r+c_m)\frac{\partial p_i}{\partial t}\right),
\end{aligned} \quad \text{(A-2)}$$

where $V_{p,i}=V_{b,i}\phi_i$, is the pore volume controlled by well i in ft^3; c_r is the rock compressibility in psi^{-1}; c_m is the fluid compressibility in psi^{-1}, for $m=o,w$, where each compressibility is

assumed to be constant; $S_{m,i}$ and $\rho_{m,i}$ in lbm/RB, respectively, are the average saturation and fluid density in $V_{p,i}$, for $m = o, w$ and ϕ_i is the average porosity of control volume i. From Fig. 2.1, we assume that, each "streamtube" which is connected to well i has half its volume within the volume controlled by well i. Then, $V_{p,i}$ is given by

$$V_{p,i} = 0.5 \sum_{j=1}^{n_{c,i}} V_{p,i,j}, \tag{A-3}$$

where $n_{c,i}$ is the number of wells connected to well i. The second integrand of Eq. A-1 denotes the mass rate of phase m flowing into the control volume i from all the connected volumes and thus can be rewritten as the sum of mass rates of all volumes connected to well i; i.e.,

$$\int_S (\rho_m \frac{k \cdot k_{rm}}{\mu_m} \frac{\partial p}{\partial n}) dS = \sum_{j=1}^{n_{c,i}} \rho_{m,i,j} A_{i,j} \frac{k_{i,j} k_{rm,i,j}}{\mu_m} \frac{\partial p}{\partial n} \mid_{i,j} , \tag{A-4}$$

where $A_{i,j}$ is the average cross-sectional area of the flow connective unit between well i and well j in ft^2; $k_{i,j}$ is the absolute permeability in the associated volume in md and $\rho_{m,i,j}$ is the phase density within the connective volume in lbm/RB. Approximating the normal derivative by

$$\frac{\partial p}{\partial n} \mid_{i,j} = \frac{p_j - p_i}{L_{i,j}}, \tag{A-5}$$

and substituting Eq. A-5 into Eq. A-4 gives

$$\int_S (\rho_m \frac{k \cdot k_{rm}}{\mu_m} \frac{\partial p}{\partial n}) dS = \sum_{j=1}^{n_{c,i}} \rho_{m,i,j} A_{i,j} \frac{k_{i,j} k_{rm,i,j}}{\mu_m} \frac{(p_j - p_i)}{L_{i,j}}. \tag{A-6}$$

Assuming that ρ_m for $m = o, w$ in the third term of Eq. A-1 is equal to $\rho_{m,i}$ yields

$$\rho_m q_{m,i} = \rho_{m,i} q_{m,i}. \tag{A-7}$$

Combining Eqs. A-2, A-6 and A-7 yields

$$\frac{1}{5.615}\rho_{m,i}V_{p,i}\left(\frac{\partial S_{m,i}}{\partial t}+S_{m,i}(c_r+c_m)\frac{\partial p_i}{\partial t}\right)$$
$$=1.127\times10^{-3}\sum_{j=1}^{n_{c,i}}\rho_{m,i,j}A_{i,j}\frac{k_{i,j}k_{rm}(S_{w,i,j})}{\mu_m}\frac{p_j-p_i}{L_{i,j}}+\rho_{m,i}q_{m,i}. \quad \text{(A-8)}$$

For $m=o,w$, we assume that the effect of fluid compressibility is sufficiently small, so that we can assume

$$\rho_{m,i}=\rho_{m,i,j},\quad j=1,2,\ldots,n_{c,i}$$

is valid. Then we can eliminate the density term in Eq. A-8 to obtain

$$\frac{1}{5.615}V_{p,i}\left(\frac{\partial S_{m,i}}{\partial t}+S_{m,i}(c_r+c_m)\frac{\partial p_i}{\partial t}\right)=1.127\times10^{-3}\sum_{j=1}^{n_{c,i}}A_{i,j}\frac{k_{i,j}k_{rm}(S_{w,i,j})}{\mu_m}\frac{p_j-p_i}{L_{i,j}}+q_{m,i}. \quad \text{(A-9)}$$

Summing Eq. A-9 for the oil and water phases gives

$$\frac{1}{5.615}V_{p,i}(c_r+S_{w,i}c_w+S_{o,i}c_o)\frac{\partial p_i}{\partial t}$$
$$=1.127\times10^{-3}\sum_{j=1}^{n_{c,i}}A_{i,j}k_{i,j}\left(\frac{k_{ro}(S_{w,i,j})}{\mu_o}+\frac{k_{rw}(S_{w,i,j})}{\mu_w}\right)\frac{p_j-p_i}{L_{i,j}}+q_{o,i}+q_{w,i}. \quad \text{(A-10)}$$

The total mobility, total compressibility and total volumetric rate, respectively, are defined by

$$\lambda_t(S_w)=\frac{k_{ro}(S_w)}{\mu_o}+\frac{k_{rw}(S_w)}{\mu_w}, \quad \text{(A-11)}$$

$$c_{t,i}=S_{o,i}c_o+S_{w,i}c_w+c_r, \quad \text{(A-12)}$$

and

$$q_{t,i}=q_{o,i}+q_{w,i}. \quad \text{(A-13)}$$

Substituting Eqs. A-11-A-13 into Eq. A-10 yields

$$\frac{1}{5.615}V_{p,i}c_{t,i}\frac{\partial p_i}{\partial t} = 1.127\times10^{-3}\sum_{j=1}^{n_{c,i}}A_{i,j}k_{i,j}\lambda_t(S_{w,i,j})\frac{p_j-p_i}{L_{i,j}} + q_{t,i}, \tag{A-14}$$

where λ_t is evaluated by upstream weighting, i.e. if $p_j > p_i$, then $S_{w,i,j} = S_{w,j}$; otherwise, $S_{w,i,j} = S_{w,i}$, where $S_{w,i}$ and $S_{w,j}$, respectively, are the saturations at the well node i and j. We define the transmissibility as

$$T_{i,j} = 1.127\times10^{-3}\frac{k_{i,j}A_{i,j}\lambda_t(S_{w,i,j})}{L_{i,j}}. \tag{A-15}$$

Substituting $A_{i,j} = \frac{V_{p,i,j}}{\phi_{i,j}L_{i,j}}$ into Eq. A-15 yields another form of $T_{i,j}$, i.e.,

$$T_{i,j} = 1.127\times10^{-3}\frac{k_{i,j}V_{p,i,j}\lambda_t(S_{w,i,j})}{\phi_{i,j}L_{i,j}^2}. \tag{A-16}$$

Substituting Eq. A-15 into Eq. A-14, we obtain

$$\frac{1}{5.615}c_{t,i}V_{p,i}\frac{\partial p_i}{\partial t} = \sum_{j=1}^{n_{c,i}}T_{i,j}(p_j-p_i) + q_{t,i}. \tag{A-17}$$

Discretizing Eq. A-17 by finite-difference in an implicit-pressure-explicit-saturation (IMPES) way gives

$$\sum_{j=1}^{n_{c,i}}T_{i,j}^{n-1}(p_j^n-p_i^n) + q_{t,i}^n = \frac{1}{5.615}\frac{c_{t,i}^{n-1}V_{p,i}^{n-1}}{\Delta t_n}(p_i^n-p_i^{n-1}), \tag{A-18}$$

where

$$T_{i,j}^{n-1} = 1.127\times10^{-3}k_{i,j}A_{i,j}\frac{\lambda_t(S_{w,i,j}^{n-1})}{L_{i,j}} = T_{i,j}^0\frac{\lambda_t(S_{w,i,j}^{n-1})}{\lambda_t(S_{w,i,j}^0)}, \tag{A-19}$$

with $\lambda_t(S_{w,i,j}^0)$ specified as the total mobility at initial water saturation and $S_{w,i,j}^{n-1}$ is equal to the saturation of the upstream well node between the connection (i,j) at time level $n-1$. By assuming that reservoir is water wetting and the initial water saturation is equal to

irreducible water saturation, S_{iw}, and that $k_{ro}(S_{iw}) = 1$, it follows that

$$\lambda_t(S^0_{w,i,j}) = \frac{1}{\mu_o}.$$

From Eq. A-3, $V^{n-1}_{p,i}$ in Eq. A-18 is given by

$$V^{n-1}_{p,i} = 0.5 \sum_{j=1}^{n_{c,i}} V^{n-1}_{p,i,j}, \tag{A-20}$$

where $V^{n-1}_{p,i,j}$ is approximated by the following first-order Taylor series expansion:

$$V^{n-1}_{p,i,j} = V^0_{p,i,j}(1 + c_r(p^{n-1}_{i,j} - p^0)), \tag{A-21}$$

where p^0 is the initial reservoir pressure, and

$$p^{n-1}_{i,j} = 0.5(p^{n-1}_i + p^{n-1}_j).$$

The expression for $c^{n-1}_{t,i}$ in Eq. A-18 is evaluated with Eq. A-12, i.e.,

$$c^{n-1}_{t,i} = S^{n-1}_{o,i} c_o + S^{n-1}_{w,i} c_w + c_r.$$

APPENDIX B

DERIVATION OF FRONT TRACKING METHOD

The Riemann problem (Juanes and Patzek, 2003) consists of finding a (weak) solution to the initial value problem:

$$\begin{aligned} &\frac{\partial S_w}{\partial t} + \eta \frac{\partial f_w}{\partial x} = 0, \\ &S_w(x,0) = \begin{cases} S_{wl} & \text{if } x < x_0 \\ S_{wr} & \text{if } x > x_0 \end{cases}, \end{aligned} \tag{B-1}$$

where η is a constant and the partial differential equation (PDE) represents a conservation law. Assuming that $S_{wl} \neq S_{wr}$, there is a single discontinuity in the initial saturation profile at x_0. The solution to the problem is a wave connecting the left (S_{wl}) and right (S_{wr}) states. For the problems considered here, we neglect gravity and capillarity, and assume the fractional flow curve is S shaped with a convex part connected to a concave part. The characteristic wave may involve three possible wave types, namely, a single rarefaction wave, a single shock wave, and a composite rarefaction-shock wave (Juanes and Patzek, 2003).

Single Shock. A shock is a discontinuous wave existing in the form of a weak solution. A single shock might develop when two characteristics intersect and the characteristic speed of the left state is greater than that of the right state, which is the necessary condition that a single shock develops. Given two states, S_{wl} and S_{wr}, the trial shock speed σ is determined by *Rankine-Hugoniot* condition:

$$\sigma_{\text{trial}} = \eta \frac{f_w(S_{wl}) - f_w(S_{wr})}{S_{wl} - S_{wr}}. \tag{B-2}$$

To guarantee the shock wave with speed of σ_{trial} is admissible, Oleinik entropy condition must be satisfied (Oleinik, 1957) for all S_w between S_{wl} and S_{wr}, i.e., we must have

$$\frac{f_w(S_w) - f_w(S_{wl})}{S_w - S_{wl}} \geq \frac{f_w(S_{wr}) - f_w(S_{wl})}{S_{wr} - S_{wl}} \geq \frac{f_w(S_{wr}) - f_w(S_w)}{S_{wr} - S_w}. \tag{B-3}$$

For our problems with only one inflection point in $f_w(S_w)$, there is an alternative entropy condition which is weaker than the condition of Eq. B-3 but also is sufficient to ensure that the solution of Eq. B-1 is a single shock, i.e., if the given S_{wl} and S_{wr} satisfy

$$f_w^{'}(S_{wl}) \geq \sigma_{\text{trial}} \geq f_w^{'}(S_{wr}), \tag{B-4}$$

Then there is a single shock solution of Eq. B-1 which is given by

$$S_w(x,t) = \begin{cases} S_{wl}, & \frac{x-x_0}{t} < \sigma_{\text{trial}} \\ S_{wr}, & \frac{x-x_0}{t} > \sigma_{\text{trial}} \end{cases}. \tag{B-5}$$

where x_0 is the location of the single discontinuity in the initial condition.

Single Rarefaction. A single rarefaction wave is a smooth function $S_w(x,t)$ that connects S_{wl} and S_{wr}. It develops when $f_w^{'}$ increases monotonically from the left state to the right state, and for an S-shaped f_w curve with a single inflection point S_{w0}, a necessary and sufficient condition for the saturation to be a rarefaction wave is given by

$$S_{wl} > S_{wr} \geq S_{w0} \quad \text{or} \quad S_{wl} < S_{wr} \leq S_{w0}, \tag{B-6}$$

where S_{w0} is the inflection point of $f_w(S_w)$. If we know the analytical form of $f_w^{''}$, it is convenient to replace Eq. B-6 by the simpler condition given by

$$f_w^{''}(S_{wl}) \cdot f_w^{''}(S_{wr}) \geq 0 \quad \text{and} \quad f_w^{'}(S_{wl}) < f_w^{'}(S_{wr}). \tag{B-7}$$

The closed form solution of a single rarefaction is given by

$$S_w(x,t) = \begin{cases} S_{wl}, & \frac{x-x_0}{t} < \eta f_w^{'}(S_{wl}) \\ (f_w^{'})^{-1}(\frac{x-x_0}{\eta t}), & \eta f_w^{'}(S_{wl}) < \frac{x-x_0}{t} < \eta f_w^{'}(S_{wr}) \\ S_{wr}, & \frac{x-x_0}{t} > \eta f_w^{'}(S_{wr}) \end{cases}. \tag{B-8}$$

Composite Rarefaction-Shock. When the left and right states are in different convexity regions of $f_w(S_w)$, the solution can involve one rarefaction wave and one shock. In this situation, a single shock and a single rarefaction connect at some intermediate point S_{w*} called the post-shock value (Juanes and Patzek, 2003). This is the value of S_w at which the left characteristic speed (rarefaction wave) equals the speed of the right discontinuity (shock), i.e.,

$$\sigma_* = \eta f_w^{'}(S_{w*}) = \eta \frac{f_w(S_{w*}) - f_w(S_{wr})}{S_{w*} - S_{wr}}, \tag{B-9}$$

where σ_* is the trial shock speed. After obtaining the post-shock value of S_{w*}, we check the Oleinik entropy condition in Eq. B-10 to see if S_{w*} to S_{wr} is an admissible shock. The shock is admissible if for all S_w between S_{wr} and S_{w*},

$$\frac{f_w(S_w) - f_w(S_{w*})}{S_w - S_{w*}} \geq \frac{f_w(S_{wr}) - f_w(S_{w*})}{S_{wr} - S_{w*}} \geq \frac{f_w(S_{wr}) - f_w(S_w)}{S_{wr} - S_w}. \tag{B-10}$$

If the condition of Eq. B-10 is violated, then the Riemann solution must not be the composite rarefaction-shock wave. Otherwise, to confirm the solution is the composite wave, we have to verify there is an admissible rarefaction wave behind the admissible shock, i.e., $f^{'}(S_w)$ is monotonically increasing from S_{wl} to S_{w*}. The closed-form of the composite rarefaction-shock is given by

$$S_w(x,t) = \begin{cases} S_{wl}, & \frac{x-x_0}{t} < \eta f_w^{'}(S_{wl}) \\ (f_w^{'})^{-1}(\frac{x-x_0}{\eta t}), & \eta f_w^{'}(S_{wl}) < \frac{x-x_0}{t} < \eta f_w^{'}(S_{w*}) \\ S_{wr}, & \frac{x-x_0}{t} > \eta f_w^{'}(S_{w*}) \end{cases}. \tag{B-11}$$

A simpler way to look at the problem is to recognize that there are only three types of possible solutions, (i) a single shock wave, (ii) a single rarefaction wave, and (iii) a composite solution consisting of a shock wave and a rarefaction wave. If we rule out (i) and (ii), the only possibility is (iii). In general, combining the three solution types, the complete algorithm to solve the Riemann problem defined in Eq. B-1 is described in the algorithm below, which represents the solution obtained by the method of characteristics (MOC) (Yuan et al., 2016).

Algorithm 1: Complete Algorithm to Solve Riemann Problems

1. Define left and right states, S_{wl}, S_{wr}.

2. Compute the trial shock speed from Rankine-Hugoniot jump condition: $\sigma_{\text{trial}} = \eta \frac{f_w(S_{wl}) - f_w(S_{wr})}{S_{wl} - S_{wr}}$.

3. IF $f_w^{'}(S_{wl}) \geq \sigma_{\text{trial}} \geq f_w^{'}(S_{wr})$ THEN

   - S: Single Shock with speed $\sigma = \sigma_{\text{trial}}$.

   ELSE IF $f_w^{'}(S_{wl}) < f_w^{'}(S_{wr})$ and $f_w^{''}(S_{wl}) f_w^{''}(S_{wr}) > 0$ THEN

   - R: Single rarefaction,

   ELSE

   - RS: Composite rarefaction-shock.
   - Compute the post-shock value S_{w*}, by solving $f_w^{'}(S_{w*}) = \frac{f_w(S_{w*}) - f_w(S_{wr})}{S_{w*} - S_{wr}}$.

   END

4. END

If we use the power-law relative permeability curve, the fractional curve has an analytical form and the first and second order derivatives can be obtained easily analytically. So the criteria to determine the solution types can be computed easily. Eq. B-9 is solved by the Newton-Raphson method.

B.1 Front Tracking Algorithm

The Riemann problem (Eq. B-1) is a simple form of a Cauchy problem with the initial condition given by a step function with a single discontinuity. In the application of the INSIM model (Zhao et al., 2015), we often encounter the Cauchy problem with nonuniform initial conditions given by

$$\begin{aligned} &\frac{\partial S_w}{\partial t} + \eta \frac{\partial f_w}{\partial x} = 0, \\ &S_w(x,0) = S_{w0}(x), \end{aligned} \tag{B-12}$$

where the initial condition is not a piecewise constant function of x. If the function $S_{w0}(x)$ is replaced by a combination of an infinite number of constant states, we could divide the Cauchy problem into an infinite number of local Riemann problems, which can be solved with the help of Algorithm 1. By connecting solutions of local Riemann problems, theoretically, we can obtain the exact solution for the problem of Eq. B-12. In practice, the initial condition should be discretized into a finite number of piecewise constant states where every two neighboring constant states define a sub-Riemann problem. By connecting the solutions of all sub-Riemann problems, Eq. B-12 is properly solved. However, if the initial saturation of Eq. B-12 is the initial condition of a classic Buckley-Leverett problem, the discretization of the initial condition is no longer needed.

The front tracking method presented here is similar to the one that was originally proposed by Holden et al. (1988) to generate the approximate solution to the Cauchy problem of Eq. B-12. In the most general situations, we need to discretize the continuous initial condition by a piecewise constant function as shown in Fig. B-1: the Cauchy problem can be decoupled into a set of local Riemann problems. Each Riemann problem shares a common constant state with its neighbor. We can use the Riemann solver given in Algorithm 1 to solve these local Riemann problems to obtain the exact MOC solutions. As shown in Fig. B-2, because the Riemann solutions which involve a rarefaction wave are spreading into a fan shape in the x-t diagram, the Riemann solutions (including the single shock case) of one local Riemann problem are also referred to as a Riemann fan (Lie and Juanes, 2005). These Riemann fans as the solutions of a series of local Riemann problems are well-defined until the

characteristics of two neighboring Riemann fans intersect each other in the $x-t$ diagram. The intersection is also referred to as a collision.

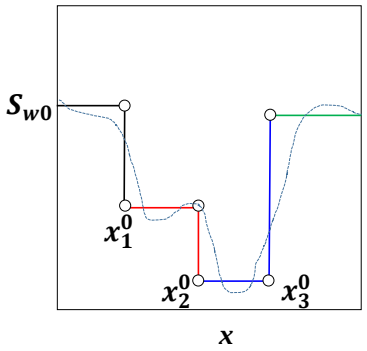


Figure B-1: Approximation of the initial condition. Here the initial condition of S_w is a continuous function of x. We approximate it by a piecewise constant function with four constant states, which produce three local Riemann problems between each two neighboring constant states.

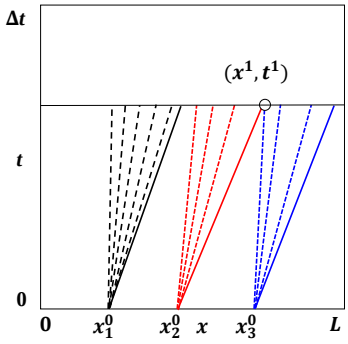


Figure B-2: Solution waves of three local Riemman problems shown in Fig. B-1. Each of the dash lines represents a characteristic in the corresponding rarefaction. Each of the solid lines represents a shock. The solutions of all these three Riemann problems are the composite rarefaction-shock waves. These three Riemann fans are valid until the shock in the second Riemann fan intersects with the first characteristic in the third Riemann fan. A new Riemann problem generates at the intersection point (x^1, t^1).

The collisions of Riemann fans often yield a curved shock path in the $x-t$ diagram, where the shock speed is changing as propagation continues. The computation of a curved shock path is computationally expensive since it is associated with numerical integration. The common way to avoid the complication of directly computing the curved shock path is to eliminate the rarefaction waves (Holden et al., 1988). In our application, the continuous

rarefaction wave in each Riemann fan is approximated with a sequence of small shocks. Each small shock is regulated to have a shock length smaller than a tolerance δS_w to guarantee numerical accuracy. The shock speed is computed with the Rankine-Hugoniot jump condition in Eq. B-2, even though the Oleinik entropy condition may be violated. Assuming one rarefaction wave in a Riemann fan ranges between S_{wL} and S_{wR}, where S_{wL} is in upstream of S_{wR}, then given δS_w, we calculate the largest integer N which satisfies

$$N \leq \frac{|(S_{wL} - S_{wR})|}{\delta S_w}.$$

Then, the interval between S_{wL} and S_{wR} is equally divided into N subintervals, which yields a vector of water saturations,

$$[S_{w0}, S_{w1}, \ldots, S_{wN}],$$

where $S_{w0} = S_{wL}$ and $S_{wN} = S_{wR}$. In this vector, every pair of two connected saturations $S_{w,j}$ and $S_{w,j+1}$ represents an approximate shock combining the saturation between $S_{w,j}$ and $S_{w,j+1}$. By doing this approximation, the solutions to all local Riemann problems are composed of a set of shocks, which are referred to as shock clusters. To adapt to the front tracking algorithm, the Riemann solver in Algorithm 1 is revised accordingly to obtain Algorithm 2.

Algorithm 2: Calculation of Shock Clusters

1. Given left and right states: S_{wl}, S_{wr}, δS_w,

2. Compute the trial shock speed from the Rankine-Hugoniot jump condition: $\sigma_{trial} = \eta \frac{f_w(S_{wl}) - f_w(S_{wr})}{S_{wl} - S_{wr}}$.

3. IF $f'_w(S_{wl}) > \sigma_{trial} > f'_w(S_{wr})$ THEN

   - S: Single Shock with speed $\sigma = \sigma_{trial}$.

ELSE IF $f_w^{'}(S_{wl}) < f_w^{'}(S_{wr})$ and $f_w^{''}(S_{wl}) f_w^{''}(S_{wr}) > 0$ THEN

- Use multiple small shocks to approximate the rarefaction wave that connects S_{wl} and S_{wr}.

ELSE

- Use multiple shocks to approximate the rarefaction wave that connects S_{wl} and S_{w*} which follows a real shock which combines the saturation between S_{w*} and S_{wr}.
- Compute the post-shock value S_{w*}, by solving $f_w^{'}(S_{w*}) = \frac{f_w(S_{w*}) - f_w(S_{wr})}{S_{w*} - S_{wr}}$.

END IF

4. END

Choosing a large value of δS_w tends to cause a large material balance error but requires less computational time to obtain solutions. Based on computational experiments, in order to balance the error and computational efficiency, we choose $\delta S_w = 0.01$. After solving all the local Riemann problems defined for the initial condition and approximating the rarefaction by small shocks, as shown in Fig. B-3, we obtain a set of shock clusters, which propagate until the first collision occurs. It is easy to see that collisions always happen at the boundary between two neighboring shock clusters. By computing all the collision times between any two neighboring shock clusters, we can find the earliest collision. For example, in Fig. B-3, there are three shock clusters initially. After computing the collision time between each two neighboring shock clusters, it is seen that the earliest collision occurs between the second shock and third shock cluster at (x^1, t^1). At the particular collision, the new Riemann problem is given by

$$\frac{\partial S_w}{\partial t} + \eta f_w^{'}(S_w)\frac{\partial S_w}{\partial x} = 0, \ t > t^1,$$
$$S_w(x, t^1) = \begin{cases} S_{wl}, & x < x^1 \\ S_{wr}, & x > x^1 \end{cases}, \tag{B-13}$$

where left (S_{wl}) and right states (S_{wr}) are given by the values immediately to the left and to the right of the colliding shocks.

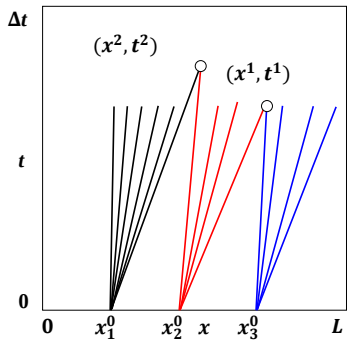


Figure B-3: A case of three shock clusters.

After solving the newly generated Riemann problem with Algorithm 2, the new shock cluster is inserted at (x^1, t^1) as represented by the dark green lines in Fig. B-4. Then the next collision is found and the associated new shock cluster is found and inserted at the collision point. The entire procedure is repeated until the earliest collision happens outside the time-space domain, $\{(x, t) \mid 0 \leq x \leq L, 0 < t < \Delta t\}$, where in our application of INSIM-FT, L is the connection length of the well pair where we need to solve for the saturation and Δt is the time step used in pressure equations. Once the loop terminates, the saturation profile at the end of the time step, Δt, can be calculated easily. The saturation profile obtained always remains piecewise constant, which is an acceptable form of the initial condition required for front tracking over the next time step. The front tracking algorithm is summarized in Algorithm 3 below.

Algorithm 3: Front Tracking Algorithm

1. Define the piecewise constant initial function,

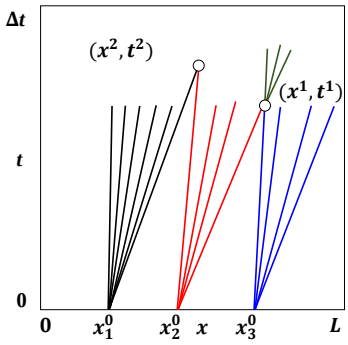


Figure B-4: Insert the new shock cluster

$$S_w^0 = \begin{cases} S_{w1}^0 & 0 < x < x_1^0 \\ \cdots & \\ S_{wi}^0 & x_{i-1}^0 < x < x_i^0 \\ \cdots & \\ S_{wn}^0 & x_{n-1}^0 < x < L \end{cases}.$$

2. Obtain solutions for all the local Riemann problems with Algorithm 2.

3. Compute the earliest collision.

4. Loop while the collision happens in $\{(x,t)|(0 < x < L, 0 < t < \Delta t)\}$,

   - Find the new Riemann solution and insert into the collision point.
   - Compute the next collision.

   End

5. Compute saturation profile at Δt and make it the new initial condition for the front tracking at next time step.

APPENDIX C

REVIEW OF CRM

C.1 Estimation of Total Production Rate

Considering the reservoir as a single tank with a single producer and a single injector, the material balance for this tank is given by

$$\frac{1}{5.615}V_b\frac{d(\phi\rho_w\overline{S}_w)}{dt} = \rho_w\left(i(t) - q_w(t)\right), \tag{C-1}$$

and

$$\frac{1}{5.615}V_b\frac{d(\phi\rho_o\overline{S}_o)}{dt} = -\rho_o q_o(t), \tag{C-2}$$

where V_b is the bulk volume of the tank in ft^3; $\overline{S}_o$ and $\overline{S}_w$, respectively, are the average oil and water saturations in V_b; ρ_o and ρ_w in lbm/RB are the oil and water densities evaluated at the average pressure $\overline{p}$ (psi) within V_b (ft^3). Under the subsurface condition, $i(t)$ represents the water injection rate of the injector in RB/day; q_o represents the oil production rate and q_w represents the water production rate of the producing well in RB/day. All rates i, q_o, q_w are nonnegative. Differentiating Eq. C-1 and C-2, respectively, and eliminating ρ_o and ρ_w gives

$$\frac{1}{5.615}V_p\left(\frac{d\overline{S}_w}{dt} + \overline{S}_w(c_r + c_w)\frac{d\overline{p}}{dt}\right) = i(t) - q_w(t), \tag{C-3}$$

and

$$\frac{1}{5.615}V_p\left(\frac{d\overline{S}_o}{dt} + \overline{S}_o(c_r + c_o)\frac{d\overline{p}}{dt}\right) = -q_o(t), \tag{C-4}$$

where $V_p = V_b \cdot \phi$; the rock and fluid compressibilities, c_r, c_w and c_o are in psi^{-1}. Summing

Eq. C-3 and Eq. C-4 yields,

$$\frac{1}{5.615}V_p c_t \frac{d\overline{p}}{dt} = i(t) - q_t(t), \tag{C-5}$$

where c_t is the total compressibility and is given by $c_t = c_o S_o + c_w S_w + c_r$; the total production rate as a function of time, $q_t(t)$, is given by $q_t(t) = q_w(t) + q_o(t)$. In CRM, c_t is constant because c_o and c_w are assumed to be constant and equal to each other. The total production rate is also written as

$$q_t(t) = J_t(\overline{p} - p_{wf}), \tag{C-6}$$

where J_t is productivity index (PI) in RB/(day $\cdot$ psi). The major versions of CRM (Yousef et al., 2005; Lake et al., 2007; Sayarpour, 2008; Weber, 2009; Nguyen, 2012; Cao et al., 2015) assume that PI is constant, i.e., does not vary with time. However, in a two-phase system, the PI of a producer is related to the saturation profile around the producer, which is generally changing with time. Thus, the assumption of constant J_t is problematic. Substituting Eq. C-6 into Eq. C-5 gives

$$\frac{1}{5.615}V_p c_t \frac{d}{dt}\left(\frac{q_t(t)}{J_t} + p_{wf}\right) = i(t) - q_t(t). \tag{C-7}$$

The time constant τ is defined by

$$\tau = \frac{c_t V_p}{5.615 J_t}. \tag{C-8}$$

Substituting Eq. C-8 into Eq. C-7 yields

$$\tau \frac{dq_t(t)}{dt} + \tau J_t \frac{dp_{wf}}{dt} = i(t) - q_t(t). \tag{C-9}$$

The general solution of Eq. C-9 is given by

$$q_t(t) = Ce^{\frac{-t}{\tau}} + e^{\frac{-t}{\tau}} \int_{\xi=t^{k-1}}^{\xi=t} e^{\frac{\xi}{\tau}} \left[\frac{1}{\tau} i(\xi) - J_t \frac{dp_{wf}}{d\xi}\right] d\xi, \quad \text{for} \quad t^{n-1} \le t \le t^n. \tag{C-10}$$

By use of the initial condition given by

$$q_t(t^{k-1}) = q_t^{k-1}, \tag{C-11}$$

the particular solution is written as

$$q_t(t) = q_t^{k-1} e^{-(\frac{t-t^{k-1}}{\tau})} + e^{-\frac{t}{\tau}} \int_{\xi=t^{k-1}}^{\xi=t} \frac{1}{\tau} e^{\frac{\xi}{\tau}} i(\xi) d\xi - e^{-\frac{t}{\tau}} \int_{\xi=t^{k-1}}^{\xi=t} J_t e^{\frac{\xi}{\tau}} \frac{dp_{wf}}{d\xi} d\xi. \tag{C-12}$$

Assuming that $i = i^k$ is constant during the time interval $\Delta t^k = t^k - t^{k-1}$, and p_{wf} is constant during the entire production history so $\frac{dp_{wf}}{dt} = 0$ yields the solution,

$$q_t^k = q_t^{k-1} e^{-(\frac{\Delta t^k}{\tau})} + i^k (1 - e^{\frac{-\Delta t^k}{\tau}}). \tag{C-13}$$

In a multi-well system of CRM, the allocation factor which is denoted by f_{ij}, is defined as the percentage of the injection rate of injector i flowing toward producer j and given as

$$f_{ij} = \frac{q_{t,i,j}(t)}{i_i(t)}, \tag{C-14}$$

where $q_{t,i,j}$ in RB/day represents the injection rate of well i flowing toward well j, while in our INSIM-FT formulation, $q_{t,i,j}$ represents the total flow rate directly from well j to well i. In CRM, f_{ij} is assumed to be constant with time even if all injection rates change, which is not strictly correct because changing injection rates will change streamlines too. The total rate flowing into the control volume of producer j at t^k is given by

$$\sum_{i=1}^{N_{\text{inj}}} i_i^k f_{ij}, \tag{C-15}$$

where N_{inj} is the number of injection wells. By analogy to Eq. C-7, the material balance

equation for $V_{p,i,j}$ is given by

$$\frac{1}{5.165}V_{p,i,j}c_t\frac{d}{dt}\left(\frac{q_{pt,i,j}(t)}{J_{t,i,j}}+p_{wf,j}(t)\right)=q_{t,i,j}(t)-q_{pt,i,j}(t), \tag{C-16}$$

where $V_{p,i,j}$ is the connection pore volume of (i,j) in ft^3; $q_{pt,i,j}$ (RB/day) is the fraction of $q_{t,j}$ (RB/day), the total liquid production rate of producer j, due to its flow connection to injector i and the volume expansion of $V_{p,i,j}$; $p_{wf,j}$ is the BHP of producer j in psi; $J_{t,i,j}$ in RB/(day $\cdot$ psi) is the productivity index of producer j that is only valid for the control volume $V_{p,i,j}$ and it follows that

$$q_{pt,i,j}(t)=J_{t,i,j}\left(\overline{p}_{i,j}(t)-p_{wf,j}(t)\right), \tag{C-17}$$

where $\overline{p}_{i,j}$ in psi is the average pressure for $V_{p,i,j}$ in ft^3. By analogy to Eq. C-9, the governing equation (Yousef et al., 2005) for well pair (i,j)

$$\tau_{ij}\frac{dq_{pt,i,j}(t)}{dt}+\tau_{ij}J_{t,i,j}\frac{dp_{wf,j}(t)}{dt}=i_i(t)f_{ij}-q_{pt,i,j}(t), \tag{C-18}$$

where τ_{ij} is defined as

$$\tau_{ij}=\frac{c_tV_{p,i,j}}{5.615J_{t,i,j}}. \tag{C-19}$$

In our applications, we consider all producers are operated at constant BHP's (or operated at varying total production rates which equivalently yield constant BHP's of producers), which yields $\frac{dp_{wf,j}}{dt}=0$ for $j=1,2,\ldots,N_{\text{pro}}$, so the solution of Eq. C-18 is given by

$$q^k_{pt,i,j}=q^{k-1}_{pt,i,j}e^{\frac{-\Delta t^k}{\tau_{ij}}}+i^k_if_{ij}\left[1-e^{\frac{-\Delta t^k}{\tau_{ij}}}\right]. \tag{C-20}$$

The total production rate for producer j is simply computed by summing $q_{pt,i,j}$, for $i=$

$1, 2, \ldots, N_{\text{inj}}$, which is given by

$$q_{t,j}^k = \sum_{i=1}^{N_{\text{inj}}} q_{pt,i,j}^k = \sum_{i=1}^{N_{\text{inj}}} q_{pt,i,j}^{k-1} e^{\frac{-\Delta t^k}{\tau_{ij}}} + \sum_{i=1}^{N_{\text{inj}}} \left[(1 - e^{\frac{-\Delta t^k}{\tau_{ij}}})(f_{ij} i_i^k) \right]. \tag{C-21}$$

Eq. C-21 is also known as the injector-producer-based representation of CRM (CRMIP) (Sayarpour, 2008), which is the version we use to compare with INSIM-FT. By matching the total production rate using Eq. C-21, optimal τ_{ij} and f_{ij} are obtained.

C.2 Estimation of Oil-Cut and Oil Production Rate with CRM

Different types of empirical models have been used to fit oil-cut. The most widely one used in a two-phase oil-water system is due to Gentil (2005) and is given by

$$f_o = \frac{1}{1 + \alpha W^{\beta}}, \tag{C-22}$$

where f_o is the estimated oil cut and W is the total volume of water injected into the reservoir and α and β are parameters to be estimated. In multiwell systems, the total volume of water injected into the control volume of producer j from time zero to time t^k is given by

$$W_j^k = \sum_{n=1}^{k} \sum_{i=1}^{N_{\text{inj}}} f_{ij} i_i^n \Delta t^n. \tag{C-23}$$

Substituting Eq. C-23 into Eq. C-22 yields

$$f_{o,j}^k = \frac{1}{1 + \alpha_j (\sum_{t=1}^{k} \sum_{i=1}^{N_{\text{inj}}} f_{ij} i_i^t)^{\beta_j}}, \tag{C-24}$$

where α_j and β_j are parameters obtained by fitting the oil cut data. This is the oil-cut model we use in CRM to compare its performance to that of INSIM-FT. Other empirical models may be better for the examples we present, but the lack of a firm physical foundation for CRM make the behaviors of these models case-dependent. Finally, the oil production rate

is estimated by

$$q_{o,j}^{k} = q_{t,j}^{k} f_{o,j}^{k}. \tag{C-25}$$

APPENDIX D

GRAHAM'S SCAN ALGORITHM

Graham's scan (De Berg et al., 2000) is an algorithm for finding the convex hull of a set X of N points in a 2D plane with time complexity of $O(n\log(n))$. The algorithm is utilized to find the upper/lower boundary or convex hull of X. The algorithm requires that the points of X are ordered as $P_1, P_2, \ldots, P_N$, where $P_1 = (x_1, y_1)$ so that $x_1 \leq x_2 \leq \ldots \leq x_N$ (or $y_1 \leq y_2 \leq \ldots \leq y_N$). Starting from a point with the minimum x coordinate (or y coordinate), the algorithm scans each point from the ordered point set X until the end to decide whether the scanned point is a valid vertex of the convex hull. Before introducing the algorithm, we need to define a function to decide if three consecutive points represent a counter-clockwise turn. Letting the function be ccw, the three given points be P_1, P_2 and P_3, respectively, and letting the x coordinate and y coordinate, respectively, of a given point P_i, $i = 1, 2, 3$, be represented by x_i and y_i, we define the function ccw by

$$\text{ccw}(P_1, P_2, P_3) = (x_2 - x_1) \cdot (y_3 - y_1) - (y_2 - y_1) \cdot (x_3 - x_1), \tag{C-26}$$

where ccw>0 means a counter-clockwise turn; ccw<0 means a clockwise turn and ccw=0 means that the three points are collinear. From a geometrical point of view, if we connect three points one by one from the point with the smallest x-coordinate to the point with the largest x-coordinate, a counter-clockwise turn means the slope of the line segment between the first two points is smaller than that of the last two points; and a clockwise turn means the slope of the line segment between the first two points is larger than that of the last two points. With the function ccw determining whether three consecutive points represent

a counter-clock wise turn or not, we can present Graham's scan algorithm to determine the upper part of the convex hull with regard to X as follows.

1. Give a set X of points in a 2D plane, sort the points by the x-coordinate, resulting in a sequence of points, $P_1, P_2, \ldots, P_N$, where $x_1 \leq x_2 \leq \ldots \leq x_N$.

2. Put the points P_1 and P_2 (with P_1 as the first point) in a list L_{upper}, which represents the list of all the vertices of the upper part of the convex hull.

3. Set $N_L = 2$ and N_L where the number of points in L_{upper}.

4. **DO** $i = 3$ to N

   - Add P_i to the end of L_{upper} and set $N_L = N_L + 1$.
   - **DO WHILE** $N_L > 2$ **AND** $\text{ccw}(P_{N_L-2}, P_{N_L-1}, P_{N_L}) \geq 0$ (where P_{N_L-2}, P_{N_L-1} and P_{N_L} represent the last three points of L_{upper}.)
     - Delete P_{N_L-1} from L_{upper} and $N_L = N_L - 1$.
   - **END DO**

   END DO

The algorithm to determine the lower part of the convex hull of X is given by replacing $\text{ccw}(P_{N_L-2}, P_{N_L-1}, P_{N_L}) \geq 0$ in the above algorithm by $\text{ccw}(P_{N_L-2}, P_{N_L-1}, P_{N_L}) \leq 0$. Note that, to find the upper part of the convex hull, we need to ensure that each three consecutive connected valid vertices of the lower part of the convex hull represent a clockwise turn (or are collinear); and vice versa to find the lower part of the convex hull.

APPENDIX E

MITCHELL'S BEST-CANDIDATE ALGORITHM

Mitchell's best-candidate algorithm (Mitchell, 1991) is an important algorithm used in the image processing area, which can sequentially generate samples that are nearly uniformly distributed in a sample space. The samples generated from Mitchell's best candidate algorithm are regulated so that the sample points are evenly distributed. The basic idea of Mitchell's best candidate algorithm is that, at each iteration, a number of candidate samples are randomly generated, and the candidate that is farthest from the existing point set is accepted as the "best candidate" for the current iteration (Dunbar and Humphreys, 2006). In our applications, the algorithm is utilized to sequentially generate imaginary well nodes within a bounded reservoir given a number of fixed actual well nodes, which can represent the perforations of actual wells. To adapt the algorithm to our applications, the original algorithm is revised to obtain the algorithm given below,

1. Specify the boundary of a reservoir, the number of actual well nodes (N_{actual}) and the coordinates of all the actual well nodes.

2. Specify the number of imaginary wells, N_I, and the number of candidates, N_p, for placing an imaginary well each iteration.

3. **DO** $i = 1, 2, \ldots, N_I$

    - Sample N_p candidate points (denoted by $C_1, C_2, \ldots, C_{N_P}$ where each point has three coordinates for 3D reservoir) for a new imaginary well Im_i within the bounded region of the reservoir from an uniform distribution. Note that these points are sampled as candidates for Im_i which is not yet placed.

- Calculate the distance from each candidate point ($C_j,\ j = 1, 2, \ldots, N_P$) to each actual well node and each imaginary well node which have been placed ($Im_1, Im_2, \ldots, Im_{i-1}$). Note that for iteration i, $i-1$ imaginary well nodes have been placed and if $i = 1$, no imaginary well nodes have been placed. For C_j, $j = 1, 2, \ldots, N_p$, the distances are denoted by a vector of $d_j = [d_{j,1}, d_{j,2}, \ldots, d_{j,N_{\text{actual}}+i-1}]$, where $N_{\text{actual}} + i - 1$ is the sum of actual well nodes and imaginary well nodes that have been placed up to iteration i.

- For $j = 1, 2, \ldots, N_p$, letting $\min(d_j)$ represent the minimum component of the vector d_j, select the point that maximizes $\min(d_j)$ over all j as the "best candidate" and place the point into the reservoir as an new imaginary well node.

END DO

To apply the algorithm, users have to specify the values of N_I and N_p. In our second example of Chapter 4, N_I is set equal to 27 and N_p is set equal to 1,000.