Re-Sampling the Ensemble Kalman Filter for Improved History Matching and
Characterizations of Non-Gaussian and Non-Linear Reservoir Models

by

Siavash Nejadi

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Petroleum Engineering

Department of Civil and Environmental Engineering
University of Alberta

Abstract

Reservoir simulation models play an important role in the production forecasting and field development planning. To enhance their predictive capabilities and capture the uncertainties in model parameters, stochastic reservoir models should be calibrated to both geologic and flow observations. The relationship between production performance and model parameters is vastly non-linear, rendering history matching process a challenging task.

The Ensemble Kalman Filter (EnKF) is a Monte-Carlo based technique for assisted history matching and real-time updating of reservoir models. EnKF works efficiently with Gaussian variables, but it often fails to honor the reference probability distribution of the model parameters where the distribution of model parameters are non-Gaussian and the system dynamics are strongly nonlinear.

In this thesis, novel sampling procedures are proposed to honor geologic information in reservoirs with non-Gaussian model parameters. The methodologies include generating multiple geological models and updating the uncertain parameters using dynamic flow responses using iterative EnKF technique.

Two new re-sampling steps are presented for characterization of multiple facies reservoirs. After certain number of assimilation steps, the updated ensemble is used to generate a new ensemble that is conditional to both the geological information and the early production data. Probability field simulation and a novel probability weighted re-sampling scheme are introduce to re-sample a new

ensemble. After the re-sampling step, iterative EnKF is again applied on the ensemble members to assimilate the remaining production history.

A new automated dynamic data integration workflow is implemented for characterization and uncertainty assessment of fracture reservoir models. This new methodology includes generating multiple discrete fracture network (DFN) models, upscaling the models for flow simulation, and updating the DFN model parameters using dynamic flow responses. The assisted history matching algorithm entails combining a probability weighted sampling with iterative EnKF.

The performances of the introduced methodologies are evaluated by performing various simulation studies for different synthetic and field case studies. The qualities of the final matching results are assessed by examining the geological realism of the updated ensemble using the reference probability distribution of the model parameters and computing the predicted dynamic data mismatch.

Dedicated to my wife, parents and sisters, for their love,

endless support and encouragement.

Acknowledgments

Table of Contents

List of Tables:

List of Figures:

List of Symbols

ψ	:	Ensemble of state vectors
ε	:	Measurement error
σ	:	Standard deviation
γ	:	Semi-variogram
Λ	:	Conditioning data
d	:	Production data
f	:	Probability distribution
g	:	Forward model
h	:	Conditioning hard data
l	:	Variogram lag distance, length of runs
m	:	Model parameters
u	:	State variables
x	:	Location
y	:	State vector
z	:	Facies indicators
C	:	Covariance matrix
F	:	Cumulative probability distribution
H	:	Observed production data operator, Heaviside function
I	:	Identity matrix
K	:	Kalman gain
N	:	Number of data
R	:	Mismatch

| Z | : | Random function of facies indicators |

Superscript

a	:	Analyzed
j	:	Ensemble member
p	:	Predicted
r	:	Re-sampled
T	:	Matrix transpose

Subscript

D	:	Measured data
b	:	Grid block
e	:	Ensemble
i	:	Index
k	:	Assimilation step
n	:	Number of lags, Maximum length of runs
w	:	Well
nf	:	Number of facies
obs	:	Observed
rp	:	Re-sampling point

Chapter 1: General Introduction

1.1 Dynamic Data Integration and History Matching

Reservoir simulation models play an important role in the production forecasting and field development planning. To enhance their predictive capabilities and capture the uncertainties in model parameters, stochastic reservoir models should be calibrated to both geologic and flow observations. Development of a practical, robust and efficient technique for dynamic data integration and history matching petroleum reservoirs is a problem of great interest. The knowledge of the parameters of the reservoir is limited to the observations at the well locations. Such data are sparsely distributed in the reservoir and data from other resources such as seismic that might be available for the whole reservoir lack precision and are indirectly related to some model parameter. Generally, lack of information casts uncertainty into the problem of reservoir characterization.

History matching is a non-linear inverse problem involving multivariate probability distributions of the reservoir model properties. The main objective of history matching is to optimize the value of uncertain static model parameters based on the dynamic observed data. History matching of continuous rock properties such as porosity and permeability distributions in heterogeneous formations can be achieved using various data assimilation techniques. An increasing effort is being made to develop organized methodologies for assessing the uncertainties in the models parameters and their conditional distributions. Advances in computer technology and the availability of low cost

and high speed computers have also generated an interest in the application of automated optimization techniques. Many authors such as Oliver and Chen (2011) have presented thorough reviews of different available techniques in recent years.

1.1.1 Gradient based method

In history matching the unknown model parameters can be estimated by minimizing an expression called the objective function. The objective function is normally based on the square difference of the predictions as compared to the observed data.

$$L(X) = \left\| f_{history}(X,t) - f(X,t) \right\|^2 ,$$

1-1

where $f_{history}(X,t)$ is the historical observation and $f(X,t)$ is the response of the model parameter state vector X.

Algorithms similar to Newton's method are applied further to update the vector of the unknown model parameters.

$$X_{k+1} = X_k + \alpha_k P_k ,$$

1-2

where k is the iteration step, α_k is the step size and P_k is the search direction. Several optimization techniques rely on the aforementioned procedure. These methods fall into the category of Gradient-Based methods. In order to find the minimum of the objective function L, its gradient with respect to X should be zero.

2

$$\nabla_X L = 0.$$
$$1\text{-}3$$

In other words, the derivative of L with respect to all elements of the model parameter vector X should be zero. Using the Taylor Series expansion, the following expression is derived:

$$X_{k+1} = X_k - H^{-1}(X_k).\nabla L(X_k),$$
$$1\text{-}4$$

where $H(X_k)$ is the Newton Hessian matrix, described by $\nabla\left(\left(\nabla L(X_k)\right)^T\right)$. One difficulty with solving reservoir history matching problems using gradient based methods is the computation of the Hessian matrix.

A major drawback of the gradient based approaches is the fact that the reference statistics of the model parameters are not taken into account. The initial guess is normally generated based on the reference statistics such as a reference distribution and a variogram (describing the spatial correlation). During the successive updates of the model parameter, the term which is added to the first guess, destructs the reference statistics.

Besides, the technique by definition, searches in the neighbourhood of the initial guess for the optimal solution. As a result, the solution is a local minimum which may be far away from the global optimum that is the scope of the history matching of petroleum reservoirs.

1.1.2 Adjoint method

As mentioned earlier, one of the drawbacks of the gradient based techniques is the difficulty in the calculation of the gradients and the Hessian matrix. In other

words, the sensitivity of the model variables such as pressure and saturation to the changes in the model parameters should be known. The gradients needed in the gradient based optimization techniques can be efficiently calculated by solving the adjoint system of equations.

The adjoint method is a well-known approach for history matching where the gradient of the loss function (squared data mismatch) and in particular the direction of the change is calculated by the adjoint method. The adjoint equation provides the sensitivity coefficients. Oliver et al. (2008) have presented some simple examples for the application of adjoint method for the history matching of a one dimensional problem. They have further described the technique in detail for complicated problems.

1.1.3 Markov Chain Monte Carlo

Markov Chain Monte Carlo has been extensively implemented in different fields to construct models of uncertain events. Its applications vary from financial and business applications to optimization problems. Markov Chain Monte Carlo is an iterative method which is based on the Markov chain that eventually converges to a fixed probability distribution. Model parameters can be considered as a Markov chain if the probability of the posteriors depends on the priors.

Each iterative step of this method consists of a proposal and an acceptance step. The model parameters of the simulation model are considered as a Markov chain. The probability of generating a realization only depends on the preceding realization (Oliver et al. 2008). Consider a realization of model parameter to be

$X^i(u)$, where i represents the step of the Markov chain, then each possible model realization has a probability π^i with $X^i(u)$. It is desired to proceed with the steps in the chain towards a realization that matches the observations with a root mean square error which is less than the previous realization. The successive update of the model parameter is carried on until an acceptable match of observed data is achieved. The method defines a transition probability Pij that should be cautiously specified for the changes from state i to state j. The probability associated with state j (π^j) is defined as the sum of the products of transition probability Pij and the probability of state i (π^i).

$$\pi^j = \sum \pi^i Pij$$

1-5

An acceptable transition matrix is determined based on several conditions. The conditions must be satisfied to make the Markov chain stationary (Goodman 1999).

Transition matrix should satisfy a reversibility condition:

$$\pi^i P_{ij} = \pi^j P_{ji}$$

1-6

The transition matrix is written as a product of two components:

$$P_{ij} = \alpha_{ij} q_{ij}$$

1-7

where q_{ij}, is the probability of proposing a transition state, and α_{ij}, is the acceptance probability.

Hasting proposed an acceptance probability in the following form:

$$\alpha_{ij} = \min\left\{1, \frac{\pi^i q_{ij}}{\pi^j q_{ji}}\right\}.$$

1-8

If the proposed transitions are symmetric, accepting a proposal is conditioned to the ratio of the probability of being in the two states. If the proposed transition is rejected, the old state is repeated in the chain.

The prescribed technique has the capability to generate realizations that are conditioned to the hard data and at the same time honor the reference statistics. In order to condition the data to the observed production history, the acceptance probability should account for the conditional probability based on the observations. The acceptance probability ratio is expressed by means of the Bayes rule that incorporates the likelihood function that is the probability of observing the dynamic data given a particular model parameter vector. After changing the model parameters, the reservoir model should be ran from the beginning to account for the state variables and estimate the likelihood.

Srinivasan et al. (2000) used proxy functions to reduce the excessive computations for the evaluation of the likelihood functions. This multipoint proxy captures the underlying non-linear relationship between the input permeability field and the output response variables using neural network theory. The multipoint proxy function significantly reduces the computational costs. However the Markov Chain Monte Carlo method itself requires large number of iterations to converge to a stationary distribution. This is mainly because of low acceptance ratios for transitions to a new state when the number of parameters in a model is large.

1.1.4 Simulated annealing

Annealing is a heat treatment process in which a solid material is heated above a critical temperature, and then cooled. The slow cooling schedule of the process causes the solid material to form a new homogeneous crystalline structure with a minimal energy. As a result of new crystalline structure the solid will have altered properties such as hardness and strength. By analogy, simulated annealing is a probabilistic method which searches for a global minimum implementing certain cooling schedules.

The idea of simulated annealing comes from Metropolis et al. (1953), where they numerically simulated the molecular behaviour when the energy level of the system is altered. Kirkpatrick et al. (1983) implemented the idea of the Metropolis algorithm and applied it to optimization problems. The idea is to use a control parameter, like temperature, to search for feasible solutions and find the global minimum of the objective function.

Simulated annealing algorithm has been applied to find the global minimum in a number of engineering problems; namely, statistical mechanics, image processing, hydrology, seismic inversion and history matching of petroleum reservoirs. It has shown potential as a global optimization technique for history matching of petroleum reservoirs. Besides, simulated annealing has the capability to generate geologically realistic model parameters using data from various resources. Hard and soft reservoir data are utilized within the algorithm to prepare solution proposals.

Despite the aforementioned statements, application of simulated annealing to petroleum reservoir characterization is too expensive. Actual reservoir models are composed of thousands of grid blocks, and the parameters of each and every grid block has to be optimized. History matching and finding the global optimum of such large model parameter space through simulated annealing requires many iterations and lots of computational efforts (e.g. Gomez et al. 2001).

Acceptance criteria

An essential part of the simulated annealing is the regulation for accepting and rejecting a certain perturbation. The acceptance criteria in simulated annealing have its route back into thermodynamics, where the probability of an increase in energy (δE) of a substance at temperature, T, is described by:

$$P(\delta E) = \exp\left(\frac{-\delta E}{kT}\right),$$

1-9

where k is the Boltzmann's constant. By analogy in simulated annealing the probability for the acceptance criteria is described by:

$$P = \exp\left(\frac{-(e_{j+1} - e_j)}{T_j}\right),$$

1-10

where T_n is a monotonically decreasing positive number derived from the cooling schedule and e_j is the magnitude of the objective function at step j.

Cooling schedule

Based on equation 1-10, at high temperatures many solutions are accepted, and as the system cools down, very few bad guesses are accepted. Initially a high temperature is used to permit movements to any neighbourhood state. However,

8

if the starting temperature is too high, it allows any movement and will become a random search. As the data are matched the temperature is reduced to admit good model proposals. The temperature can decrease until it reaches zero, or a suitably low temperature or when the system is unchanging at the any temperature.

Discretized space

While implementing simulated annealing algorithm, same as other global optimizations techniques, the space of the estimation parameters should be discretized. In different problems the model parameters might take value from continues distribution. However, practically it requires excessively large computational time to optimize a value from a continuous space. As a result the parameter estimation space is discretized into finite number of discrete values.

$$\Delta U_i = \frac{U_i^{max} - U_i^{min}}{N},$$

1-11

1.1.5 Kalman filter algorithms

Kalman filter

The Kalman filter (Kalman 1960) is a set of mathematical equations which uses a series of measurements observed over time to estimate the state of a process. It is a predictor corrector (recursive) algorithm which minimizes the mean square error. An underlying assumption of the Kalman filter is that the system dynamics are linear and the errors and measurements follow a Gaussian distribution.

The Kalman filter deals with the problem of estimating the current state of a process that is described by the following stochastic differential equation:

9

$$x_k = A_k x_{k-1} + B_k u_k + w_{k-1},$$

1-12

where, A_k is the state transition model, B_k is the control input model that is applied to the control vector u_k and w_{k-1} is the process noise with a mean of zero and a covariance of Q :

$$P(w) \sim N(0,Q).$$

1-13

The measurement z is defined as:

$$z_k = H_k x_k + v_k.$$

1-14

where, H_k is the observation model and v_k is the measurement noise with a mean of zero and a covariance of R :

$$P(v) \sim N(0,R).$$

1-15

The priori state estimate, which is based on the knowledge of the system at step k, is defined as $\hat{x}_k^-$. The posterior state estimate at step k, using the measurements z_k is defined as $\hat{x}_k$. The prior estimate error is defined as $e_k^- = x_k - \hat{x}_k^-$ and the posterior error is $e_k = x_k - \hat{x}_k$.

The correctness of the state estimates is measured by means of error covariance matrices. The priori estimate error covariance is defined as:

$$P_k^- = E\left[e_k^- e_k^{-T}\right].$$

1-16

The posteriori estimate error covariance is defined as:

$$P_k = E\left[e_k e_k{}^T\right].$$

1-17

The posteriori state estimation $\hat{x}_k$ is defined as a linear combination of the priori estimate $\hat{x}_k^-$ and a weighted difference between an actual measurement and the predicted measurement:

$$\hat{x}_k = \hat{x}_k^- + K\left(z_k - H\hat{x}_k^-\right),$$

1-18

where K is the Kalman gain and minimizes the posteriori noise covariance:

$$K = \left(P_k^- H^T\right)\!/\!\left(H P_k^- H^T + R\right)^{-1}.$$

1-19

The Kalman filter can be described as two different steps, predict and update. The prediction phase uses the current state to estimate the state at the next step. In the update phase, the priori estimates and the current observations are used to improve the state estimate.

As mentioned earlier, Kalman filter minimizes the mean square error. The expected value of the square error in the posteriori estimate error is:

$$E\left[(x_k - \hat{x}_k)^2\right].$$

1-20

Minimizing the expected value of the square error is the same to, minimizing the trace of the posteriori estimate error covariance matrix, which is P_k. The trace is minimized when its derivative with respect to the gain is set to zero. This leads to the Kalman gain which minimizes the mean square error.

Extended Kalman filter

The extended Kalman filter is an extension of the Kalman filter, to address the filtering problem when the system dynamics are nonlinear. For nonlinear problems, the relationship of the state and observations can be written as:

$$x_k = f(x_{k-1}, u_k) + w_{k-1}.$$

1-21

The data relationships are also defined as:

$$z_k = g_k(x_k) + v_k.$$

1-22

Predict and update steps of the extended Kalman filter are similar to Kalman filter. The terms in the formulations are defined as follows:

$$A_k = \left.\frac{\partial f}{\partial x}\right|_{\hat{x}_{k-1}, w_{k-1}}.$$

1-23

$$H_k = \left.\frac{\partial g}{\partial x}\right|_{\hat{x}_k}.$$

1-24

The formulation means that the nonlinear relationships are approximated by linearized relations. In other words, extended Kalman filter linearizes about the mean and the covariance. As a result, if the predict and the update functions are highly nonlinear, application of the extended Kalman filter will result in poor estimates. Furthermore, it is noteworthy that, application of extended Kalman filter requires computation of the H_k, which is cumbersome and computationally expensive.

Unscented Kalman filter

The Unscented Kalman filter is used to address the approximation issues of the extended Kalman filter. It uses unscented transform to pick a set of sample points around the mean. As a result it is a derivative free technique and also provides superior estimates with a same computational efficiency as extended Kalman filter. Unscented transform gives correct mean up to third order and covariance up to the second order for any function. Linearized mean is correct only up to first order. Covariance in unscented transform and linearization has the same order of accuracy. However, the magnitude of the error is smaller in unscented transform.

1.2 Ensemble Kalman Filter

The Ensemble Kalman Filter (EnKF) has gained popularity over recent years as a Monte-Carlo based technique for assisted history matching and real time updating of reservoir models (Lorentzen et al. 2001; Aanonsen et al. 2009). The EnKF procedure utilizes an ensemble of model states (e.g. realizations of reservoir properties such as porosity and permeability) to approximate the covariance matrices used in the updating process. The initial ensembles are generated based on prior knowledge of the reservoir, while the sequential updates lead to a sampling of the posterior probability function.

Recently, EnKF has been successfully used for history matching of simple conventional reservoirs (a complete list is available in Oliver et al. 2011). However, it has the potential to be implemented for history matching and characterization of unconventional reservoirs with complex processes. Such

applications range from reservoir characterization of heavy oil reservoirs that undergo thermal recovery processes such as SAGD, to fractured reservoirs with complex geology and fluid flow patterns (Gul et al. 2011, Nejadi et al. 2012a). EnKF propagates an ensemble of initial reservoir realizations along time to assimilate data.

The main procedure for EnKF contains two parts. Firstly, in the forecast step, the forecast model is applied to each ensemble separately, using the reservoir simulator:

$$\begin{bmatrix} u_k^{p,j} \\ d_k^{p,j} \end{bmatrix} = g \begin{bmatrix} m_{k-1}^{a,j} \\ u_{k-1}^{a,j} \end{bmatrix}, \quad j = 1,...,N_e ,$$

1-25

In the above equation, m and u denote the model and state variables, respectively, and d_k represents the predicted production data at k^{th} step. The superscripts, p and a, specify the predicted and analyzed states, respectively.

Fluid flow in porous media is governed by the law of conservation of mass and Darcy's equation. Multiphase fluid flow in a reservoir with no flow boundary is described by the following partial differential equation:

$$\nabla.\left(\alpha\rho_l v_l\right) + \alpha \frac{\partial\left(\rho_l\phi s_l\right)}{\partial t} = \alpha\rho_l q_l .$$

1-26

In the above equation l represents the fluid phase that is oil and water $l = \{o,w\}$. The velocity term v_l of phase l is calculated from Darcy's law

$$v_l = -\frac{Kk_{rl}}{\mu_l}\left(\nabla p_l - \rho_l g\nabla d\right),$$

1-27

14

where p_l and μ_l denote the pressure and viscosity of phase l respectively. g is the gravitational acceleration and K is the permeability tensor.

The permeability tensor, K, is in the following form:

$$K = \begin{bmatrix} k_{xx} & k_{xy} & k_{xz} \\ k_{yx} & k_{yy} & k_{yz} \\ k_{zx} & k_{zy} & k_{zz} \end{bmatrix}$$

In conventional homogenous reservoirs, the grid blocks of the reservoir simulation model are defined along three directions that transform the permeability tensor into a diagonal matrix. Specifically, a pressure drop along the coordinate directions would only yield to fluid flow in same path.

$$K = \begin{bmatrix} k_x & 0 & 0 \\ 0 & k_y & 0 \\ 0 & 0 & k_z \end{bmatrix}$$

Assuming an isotropic permeability and substituting Darcy velocity from equation 1-27 into equation 1-26 the following partial differential equation is obtained:

$$\nabla \cdot \left(-\alpha \rho_l \frac{k k_{rl}}{\mu_l} \left(\nabla p_l - \rho_l \, g \nabla d \right) \right) + \alpha \frac{\partial \left(\rho_l \phi s_l \right)}{\partial t} = \alpha \rho_l q_l .$$

1-28

Considering the following relations

$$s_o + s_w = 1 .$$

1-29

$$p_o - p_w = pc_{ow}(s_w) .$$

1-30

15

where pc denotes capillary pressure. Assuming that density depends only on pressure, the second term in equation 1-26 can be written as

$$\frac{\partial(\rho_l \phi s_l)}{\partial t} = \frac{\partial \rho_l}{\partial p_l}\frac{\partial p_l}{\partial t}\phi s_l + \frac{\partial \phi}{\partial p_l}\frac{\partial p_l}{\partial t}\rho_l s_l + \rho_l \phi \frac{\partial s_l}{\partial t} .$$

1-31

Assuming that oil and water density depends on the pressure, the compressibility terms are defined as

$$c_l = \frac{1}{\rho_l}\frac{\partial \rho_l}{\partial p_l} .$$

1-32

The rock compressibility is also defined as

$$c_r = \frac{1}{\phi}\frac{\partial \phi}{\partial p_l} .$$

1-33

Substituting equation 1-29 and the compressibility terms (equations 1-32 and 1-33) in equation 1-38, the following equations are obtained for oil and water phases:

$$-\nabla.\left(\alpha \rho_o \frac{kk_{ro}}{\mu_o}[\nabla p_o - \rho_o g \nabla d]\right)$$
$$+\alpha \rho_o \phi\left[(1-s_w)(c_o + c_r)\frac{\partial p_o}{\partial t} + \frac{\partial s_w}{\partial t} - \alpha \rho_o q_o\right] = 0$$

1-34

$$-\nabla.\left(\alpha \rho_o \frac{kk_{ro}}{\mu_o}[\nabla p_o - \rho_o g \nabla d]\right)$$
$$+\alpha \rho_o \phi\left[(1-s_w)(c_o + c_r)\frac{\partial p_o}{\partial t} + \frac{\partial s_w}{\partial t} - \alpha \rho_o q_o\right] = 0$$

1-35

The above equations are discretized in the grid system of the reservoir and considering phase behavior equations of the reservoir fluids, and the relative permeability correlations, reservoir simulation software are used to solve the

fluid flow equations which are considered as forecast model. In this study the commercial reservoir simulation software Eclipse black-oil (2011) and CMG (2013) are used as the forecast model.

In the second step of EnKF, update step is applied using the observation data (d_k) for all ensemble members. The ensemble of vectors is denoted by:

$$\psi_k = \left\{ y_k^1, y_k^2, \ldots y_k^{N_e} \right\},$$ 1-36

where N_e is the total number of realizations and $y^i, i = 1, \ldots, N_e$ are state vectors. State vector consist of model parameters (m), such as porosity and permeability, state variables (u), such as pressure and saturation, which are time dependent variables, and observations (d), such as production and pressure records.

$$y_k = \left\{ \begin{array}{c} m_k \\ u_k \\ d_k \end{array} \right\},$$ 1-37

where k denotes the time step in which data are assimilated.

In the second (analysis) step, the state vectors are updated with the observation data d_k^j using following equation

$$y_k^{a,j} = y_k^{p,j} + K_k \left(d_k^j - H y_k^{p,j} \right),$$ 1-38

The forecast and analysis steps are repeated sequentially in time whenever new observations are available. H is the observation operator, which represents the relationship between the state vector and the observation vector

$$H = [0 \mid I],$$

<div align="right">1-39</div>

where d_k^j is the summation of the observed production data at the kth assimilation step $(d_{obs,k})$ and ε_k^j, a vector of measurement errors such that

$$E[\varepsilon \varepsilon^T] = C_D$$

$$d_k^j = d_{obs,k} + \varepsilon_k^j,$$

<div align="right">1-40</div>

Kalman gain K_k is defined as:

$$K_k = C_{y_k^p} H^T \left(H C_{y_k^p} H^T + C_{Dk} \right)^{-1},$$

<div align="right">1-41</div>

where $C_{y_k^p}$ is the state cross covariance matrix and C_{Dk} is the measurement error covariance matrix. The cross covariance matrix is estimated from the ensemble of state vectors as

$$C_{y_k^p} \approx \frac{1}{N_e - 1} \left\{ \left(y_k^p - \overline{y_k^p} \right) \left(y_k^p - \overline{y_k^p} \right)^T \right\},$$

<div align="right">1-42</div>

where T indicates matrix transpose, and $\overline{y_k^p}$ is the average of all posterior state vectors. The forecast and update steps are repeated as new observations become available.

Storing large covariance matrices requires lots of memory and calculating the inverse of the covariance matrices requires large computational effort, especially when a reservoir with large number of grid block and large number of wells is tackled. While applying EnKF, in order to estimate the Kalman gain, the inverse of a matrix with a size of n by n (n=total number of observations) should be calculated. Number of observations is normally very small as compared to the

<div align="right">18</div>

number of grid blocks, as a result, calculating the inverse matrix is fast and does not require large computational effort. The $C_{y_k^{p,j}}$ matrix is fairly large (*m+n* by *m+n*, where *m* is the total number of unknown model parameters and n is the total number of observations). However, this matrix is not directly used in the computations (compare with other history matching algorithms, such as gradient based technique, where, it is required to calculate the Hessian matrix and its inverse). Using simple mathematical simplifications, calculation and storing this large matrix can be avoided.

1.2.1 Ensemble size

The Ensemble Kalman filter is highly dependent on the size and the characteristics of the initial ensemble (Houtekamer and Mitchell 2000). The ensemble should be statistically representative of the problem. In other words the ensemble should provide sufficient information to map the uncertainties associated with the model. The ensemble size is normally smaller than the size of state in EnKF. If the ensemble size is so small, it is not representative of the model and the problem is under sampled.

Practically it is not possible to include very large number of realizations. Considering the computational costs, the ensemble size should be limited for practical applications. This is of high importance for petroleum reservoir history matching problems, where simulation of each ensemble member would take several hours or days. Generally, realizations should be conditioned to the hard data and follow the reference distributions and the ensemble should be large

19

enough to incorporate the noise and the uncertainties in distribution and spatial correlation of the data. The ensembles having on the order of 100 realizations are considered to be large enough for practical applications (Houtekamer and Mitchell 2000).

Spurious correlation and non-physical updates

EnKF analysis can result in updated variables that exceed certain bounds. The issue has been reported by several authors while applying EnKF for assimilation of different physical models. Several authors have reported and discussed non-physical saturation and pressure values in application of EnKF for petroleum reservoir history matching problems (Zafari and Reynolds 2007). The saturation values should lie in the range of irreducible water saturation (S_{wi}) and 1-S_{or} in waterflooding models. However, close to the displacement front of waterflooding problems, saturation data are non-Gaussian (bimodal) and cannot be accurately approximated by a Gaussian distribution. As a result, non-physical saturation values are obtained. Normally, the saturation values beyond the physical bounds are truncated before advancing the simulator to the next time step. Even with truncation, the values are not consistent with the model variables. Lorentzen et al. (2012) have rerun the models from the initial condition at each assimilation time step to calculate the dynamic variables. They have mentioned that for their synthetic two dimensional models there is no major change in computational time. Gu and Oliver (2006) have proposed different methods (normal score transform of saturation, use of the location of water

20

shock front instead of saturations and iterative schemes) to prevent non-physical water saturations.

The limited ensemble size can also cause spurious correlations. Naturally parameters of the grid blocks which are located far from a specific observation are not correlated. However, it can introduce noise in the analysis. A common approach is to use localization techniques to prevent spurious long distance correlations and limit the influence of observed data. The technique modifies the error covariance matrices and suppresses the effect of far observations. It increases the degrees of freedom of the problem and it prevents or at least delays filter divergence (Emerick and Reynolds 2010).

Ensemble collapse and loss of ensemble variance

EnKF relies on statistical measures of the ensemble when computing the updates. If the ensemble size is not sufficiently large, sampling errors affect statistical measures and results in poor approximation to the cross covariance matrix. This leads to unphysical updates of reservoir properties and loss of ensemble variability.

For field case applications, the size of the ensemble should be kept small for computational efficiency. In such applications, poor cross covariance approximation results in very large Kalman gain values and unrealistic updates, as a result ensemble variance would diminish after several assimilation steps. The updated models collapse toward a particular response which does not honor static information and reference statistics (e.g. histogram and semi-variogram or covariance) inferred from the reference geologic information. This problem is

21

referred to as ensemble collapse and is resolved by implementing covariance localization (Zhang and Oliver 2010) and re-sampling (Emerick and Reynolds 2012, Nejadi et al. 2014 a,b).

1.2.2 EnKF for non-Gaussian reservoir model parameters

As mentioned earlier, the underlying assumption of Ensemble Kalman filter is that the prior joint pdf is Gaussian when computing the updates and EnKF will not converge to the correct distribution, identified by geological studies, if the prior joint pdf has non-Gaussian contributions (e.g., multimodal distribution, curvilinear, and channelized features). In many cases, statistics of the final updated model variables would deviate significantly from the non-Gaussian distribution exhibited by the initial model states. In other words the estimated posterior pdf is not consistent with the prior distribution (Zafari and Reynolds 2007).

This assumption of Gaussianity implies that the conventional EnKF method must be modified such that it can be applied for models whose petrophysical properties are not characterized by multivariate Gaussian distributions (Evensen 2007, Aanonsen et al. 2009).

1.2.3 Parameterization techniques

It is well-known that large scale inverse problems such as history matching of petroleum reservoirs are ill-posed and the problem does not have a unique solution. Conceptually there exists infinite number of models that match the

observed data. Parameterization techniques are applied in order to reduce the number of unknown model parameters and make the problem better-posed. Parameterization techniques have been introduced to reduce the number of model parameters and the associated computational cost to calculate the sensitivity coefficients at every location in the reservoir. A small set of critical model parameters or the basis functions of the transformed model parameters that incorporate the main portion of the model variance are retained and adjusted to history match the model. Furthermore, regularization by parameterization is implemented to transform the non-Gaussian model parameters into one or a combination of Gaussian variables.

Discrete Cosine transform

Discrete Cosine Transform (DCT) has its roots in image processing (Jain 1988, Rao and Yip 1990). It is a Fourier-based transform that is applied to decompose the model parameters into a set of orthonormal cosine functions. The coefficients of the retained cosine functions are used instead of the actual model parameters. In order to reduce the size of parameter space, largest DCT coefficients are selected and used. DCT is widely used in signal and image processing for data compression. The one-dimensional DCT (X_k) of a signal x_n of length N is defined as:

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}(n+1/2)k\right] \qquad k = 0,...,N-1.$$

1-43

The inverse DCT is defined as:

$$x_n = \sum_{k=0}^{N-1} X_k \cos\left[\frac{\pi}{N}(n+1/2)k\right] \qquad n = 0,...,N-1.$$

<div align="right">1-44</div>

From the equations, it can be seen that both transformation and the inverse are separable. In other words, a multidimensional transformation can be performed by several one dimensional transforms.

DCT is implemented to parameterize the non-Gaussian model parameters into coefficients of the retained cosine basis functions. These coefficients are incorporated into the state vector and updated in the EnKF procedure. Jafarpour et al. (2008) have introduced a DCT-EnKF procedure, where, the coefficients of the retained cosine functions were included in the state vector instead of actual model parameters. In order to reduce the size of the model parameter space and capture the spatial continuity of different facies, few DCT coefficients are included in the state vectors. If the history-matched models contain artificial short-scale variability, smoothing algorithms should be implemented (Nejadi et al. 2012b).

Wavelet transform

Similar to discrete cosine transform, wavelet transform has been extensively used in image processing algorithms. Wavelet transform provides a time frequency version of the signal (data). Wavelet transform has an advantage over the Fourier transform. Wavelet Transform implements a multi resolution technique by which different frequencies are analyzed with different resolutions. While Fourier transform gives a constant resolution at all frequencies.

$$X_{WT}(\tau,s) = \frac{1}{\sqrt{|s|}} \int x(t).\psi\left(\frac{t-\tau}{s}\right)dt$$

1-45

where $x(t)$ is the signal to be analyzed, is the basis function and is the translation parameter.

Sahni and Horne (2005) have proposed a multi-resolution wavelet analysis to integrate history data together with the geostatistical information. They have used Haar wavelets to parameterize permeability and included both dynamic data and variogram of spatial permeability distribution in the objective function. Discrete wavelet transform is a wavelet transform of discrete wavelets. Same as DCT, the main coefficients of the data are retained and the rest are set to zero. The updated (optimized) coefficients are further used to reconstruct the data (model parameters). The transformed variables of DWT incorporate the spatial distribution, while DCT parameters have no specific spatial relationship. Retained DWT parameters follow a distribution which is close to the original distribution, while DCT parameters are nearly Gaussian.

Pilot point method

The use of pilot points has been proposed by de Marsily et al. (1984) as a re-parameterization technique in traditional history matching to reduce the size of the parameter space (and the associated computational costs) during model updating. In this method, the initial realizations of model parameter distributions ($K(u)|$, for all $u \in$ reservoir) are generated through stochastic simulation conditioned to the hard data at well locations. Other reservoir information such

as soft data (seismic), geological properties and geostatistical parameters (variogram model $\gamma(h)$) are also utilized in the stochastic simulation.

The parameter estimations at the specific pilot point locations are perturbed and optimized within optimization algorithms. The changes made at the pilot points are propagated to other locations by performing conditional simulation using model values at pilot points as part of the conditioning data, together with geostatistical parameters (e.g. reference distribution and semi-variogram models), to preserve the spatial correlation of the model parameters.

A practical consideration is the specification of the location and number of the pilot points, which can be specified in advance. RamaRoa et al. (1995) proposed placing the pilot points in high sensitivity zones where they have the highest potential to reduce the objective function and correctly locate the heterogeneities. Unfortunately, the number of pilot points is inherently empirical.

The number of pilot points should be selected depending on the nature of the problem. Factors such as reservoir model, well pattern, production mechanism and reservoir heterogeneity in both vertical and horizontal directions should be taken into consideration to determine the number of pilot points. The number of pilot points may be optimized such that they are large enough to capture the adjustments made to model parameters during the optimization steps while being small enough to ensure that the entire model parameter space is sufficiently sampled.

In the pilot point formulations, the locations are selected in prior. All locations in the reservoir are sorted, based the impact of the model parameter at that specific location on the objective function. The pilot points are located in the high sensitivity zones and the estimated values at these locations are propagated to other locations, by means of conditional simulation.

1.2.4 Ensemble Kalman filter for highly non-linear history matching problems

EnKF is sequential updating data assimilation technique. The observed data are implemented to update the model parameters as they become available, without re-running the simulator from the beginning. Other history matching techniques require re-running the simulator from the initial reservoir conditions, as soon as the model parameters are updated. EnKF provides estimations of state variables, as well as the updated model parameters, at each assimilation step. Burgers et al. (1998) have shown that the updated ensemble obtained with EnKF approximates the theoretical posterior distribution for large ensemble size for (1) linear dynamics and measurements and (2) Gaussian prior and likelihood.

Applying the updating schemes iteratively during each assimilation step has been proposed to address the issues related to non-linearity (Chen and Oliver 2012; Li and Reynolds 2009). The highly nonlinear relationship between the model parameters and the simulation model production data implies that assimilating state variables (e.g. pressure and saturation) is not practical and each

proposed state requires a run of forward model after every update step to ensure consistency among state and model variables.

1.2.5 Characterization and history matching of multiple facies reservoirs

Facies modeling is a crucial component of geostatistical reservoir characterization which facilitates construction of models for complex reservoirs. Facies are commonly represented by indicators, which are intrinsically non-Gaussian. It is important to condition models to reference geologic information and dynamic flow data in reservoir characterization. In many cases, static data (e.g. conceptual models, log, core, seismic interpretations, and statistics from similar fields) are available for reservoir modeling. Therefore, characterization of a robust description of geological features such as facies distributions (proportions and spatial patterns) and channel properties (e.g. orientation, widths and amplitudes) should incorporate all the static information and honor the reference statistics (e.g. histogram and semi-variogram or covariance) inferred from the reference geologic information.

Different parameterization approaches can also be found in the literature, which are combined with EnKF for history matching of non-Gaussian reservoir model parameters. Truncated Pluri-Gaussian (Liu and Oliver 2005, Agbalaka and Oliver 2008), Level Set Method (Moreno and Aanonsen 2011), Gaussian Mixture Models (Dovera and Della Rossa 2010) and Discrete Cosine Transform (Jafarpour et al. 2008 and Nejadi et al. 2014a, b) are some of these

methodologies. Gu and Oliver (2007) implemented Normal Score transform to avoid nonphysical saturation values while updating the saturation values by means of EnKF. Li et al. (2012) and Nejadi et al. (2012a) have applied normal score transform to estimate bimodal distributions of aquifer model parameters and permeability tensors in fractured reservoirs, respectively.

1.2.6 Characterization and history matching of naturally fracture and hydraulic fractured reservoirs

Reservoir simulation models play an important role in the production forecasting and field development planning. To enhance their predictive capabilities and capture the uncertainties in model parameters, stochastic reservoir models should be calibrated to both geologic and flow observations.

Fluid flow in fractured reservoirs mainly takes place through the network of interconnected fractures surrounding matrix blocks. Proper characterization of the fracture system is of outmost importance while making a robust model for simulation of the multiphase fluid flow in fractured reservoirs. The fracture system can be characterized by probability distributions of fracture properties in a discrete fracture network model. The relationship between production performance and the fracture parameters is vastly non-linear, rendering the process of adjusting model parameters to match both static geologic and the dynamic production data challenging.

Traditionally, fractured reservoirs have been treated the same as conventional reservoirs, and dynamic model parameters such as permeability, porosity and

matrix fracture interaction coefficient are adjusted to match the field production performance.

Several authors have used the parameters of a volumetric grid, which is derived from continuous fracture models, as a tuning parameter. Ouenes et al. (1995) implemented a neural network to correlate geological information and well performance to the reservoir fracture intensity. Ouenes (2000) implemented a fuzzy neural network to evaluate the effect of different fracture drivers, such as structure, lithology and bed thickness on fractures and develop correlations between geological drivers and fracture intensity. The optimized fracture intensity map is used for selecting potential infill drilling well locations with estimated ultimate recovery higher than certain economic limits. Sezuki et al. (2007) have used probability perturbation methods to adjust fracture intensity and large-scale fracture trend. They have calculated directional effective permeability of the single porosity reservoir simulation model from matrix permeability and fracture intensity. Cui and Kelkar (2005) have used a gradient simulator and adjoint method for conditioning the fracture intensity to the production data. They have used the fracture intensity map to estimate directional fracture permeability and matrix fracture interaction coefficient.

Few authors have focused on characterizing parameters of discrete fractures integrating dynamic production data. Gang and Kelkar (2008) have used the adjoint method to calibrate permeability of individual fractures as well as the capillary pressure curves. They have assumed that the relationship between the

grid block effective permeability and fracture permeability is known. Hu and Jenni (2005) have used object-based Boolean simulation to define faults and fractures in the model. They have implemented the gradual deformation method to calibrate the object based model (location, shape and size of the objects) to dynamic data. De Lima et al. (2012) have used gradual deformation to characterize fault density, fault position and length of a fractal fault model.

1.3 Problem Statement

Developing an efficient reservoir characterization, history matching and uncertainty assessment technique is a problem of great interest.

- Ensemble Kalman filter is a Monte-Carlo based technique for assisted history matching and real time updating of reservoir models. EnKF works efficiently with multivariate Gaussian variables and linear dynamics, but it often fails to honor the reference probability distribution of the model parameters and to achieve an acceptable production data match where the system dynamics are strongly nonlinear, especially of the type related to multiphase flow, or if non-Gaussian prior models are used.

- History matching of non-Gaussian multiple facies reservoirs fails to honor the reference probability distribution of the model parameters inferred from geologic information. This increases the prediction errors of the history matched models.

Estimation of fracture parameters is often challenging because reservoir characterization based on both static and dynamic data is an inverse problem that is highly nonlinear and the solutions are not unique.

- A robust, practical technique that assimilates fracture parameters and generates multiple history matched models has not been developed for history matching and characterization of fractured reservoirs.

1.4 Research Objectives

In this research, dynamic production data and geologic information are integrated in an ensemble-based history matching technique to assimilate various model parameters of the reservoir model. The principal objective of this work is to develop practical ensemble based history matching methodologies that update reservoir models honoring both static geological data and dynamic information.

- Two new re-sampling techniques are proposed for inclusion in the conventional EnKF algorithm for characterization and history matching of non-Gaussian multiple facies reservoirs. The objectives are to (1) preserve facies proportions, (2) preserve spatial pattern and distribution of different facies, (3) maintain diversity among realizations during updating and avoid ensemble collapse.

- A novel methodology is proposed for characterization and history matching of natural fracture reservoirs and hydraulic fractured wells. Dynamic data are integrated to characterize fractures and reduce uncertainties in fracture parameters and their spatial distribution.

Multiple discrete fracture network models and their equivalent upscaled flow simulation models are generated. The models honor geological data and match the dynamic production history.

- Implementing the methodology for real history matching problems is not overlooked in this study. The new methodology is applied for characterization and history matching of a multi-stage hydraulic fractured shale gas well in the Horn River basin to show that the method can be applied to real field applications.

1.5 Thesis Layout

The rest of this thesis is organized as follows. Chapter 2 introduces a new step for inclusion in the history matching of multiple facies reservoir models using EnKF. The new step consists of constructing a facies probability map and application of probability field (P-Field) simulation to re-sample a new ensemble. Chapter 3 presents a novel probability weighted re-sampling scheme within the conventional framework of EnKF. The assimilated ensemble members are implemented to sample an improved ensemble, which incorporates the dynamic model-updating information of the initial steps and is consistent with the static geological data. In Chapter 4, an integrated approach for history matching and characterization of natural fractures is presented. This new methodology includes generating multiple discrete fracture models, upscaling them for numerical multiphase flow simulation, and updating the fracture properties using dynamic flow responses such as continuous rate and pressure measurements. Chapter 5 implements the methodology discussed in chapter 4,

for characterization of a hydraulically-fractured shale gas well in the Horn River basin. It demonstrates the applicability of the proposed framework for uncertainty quantification of hydraulic fracture parameters for shale gas reservoirs. Chapter 6 summarizes the ideas presented in this thesis along with suggestions for future research on this topic.

Chapter 2: Estimation of Facies Boundaries Using

Categorical Indicators with P-Field Simulation and

Ensemble Kalman Filter[1]

2.1 Introduction

The Ensemble Kalman Filter (EnKF) is a Bayesian updating scheme that implements the Monte-Carlo technique for data assimilation. It has been introduced to the petroleum engineering industry by Lorentzen et al. (2001) as a promising approach for solving high-dimensional history matching problems (Aanonsen et al. 2009). The EnKF estimates an ensemble of model states by approximating the covariance matrices sequentially in time as new observations become available. Each member of the ensemble is updated using an ensemble approximation to the Kalman gain, which is approximated from the mean and covariance of the prior joint probability density function (pdf). The updated ensemble provides an empirical estimate of the posterior joint probability distribution. The Kalman filter performs well when there is a linear relationship between state variables, model parameters, and the data, with the underlying assumption that the prior joint pdf being Gaussian. In theory, the EnKF does not converge to the correct distribution, identified by geological studies, if the prior joint pdf has non-Gaussian contributions. This could create inconsistencies

35

between the estimated posterior pdf and the prior distribution (Zafari and Reynolds 2007).

The conventional EnKF method must be modified such that it can be applied for models whose petrophysical properties are not characterized by multivariate Gaussian distributions (Evensen 2007; Aanonsen et al. 2009). For instance, facies modeling is a crucial component of geostatistical reservoir characterization, which facilitates the construction of models for complex reservoirs. Facies are commonly represented by categorical data, which are intrinsically non-Gaussian. Geostatistical modeling techniques have been implemented by various authors to build detailed geological models integrating static data (e.g. logs, core, seismic interpretations and statistics from similar fields). Modis and Sideri (2013), Yamamoto et al. (2012) and Teng and Koike (2007) are some of the recent studies that have demonstrated successful applications of the Truncated Gaussian and spline-based interpolation techniques. Different parameterization approaches can also be found in the literature, which are combined with EnKF for history matching of non-Gaussian reservoir model parameters. Truncated Pluri-Gaussian (Liu and Oliver 2005; Agbalaka and Oliver 2008), Level Set Method (Moreno and Aanonsen 2011), Gaussian Mixture Models (Dovera and Della Rossa 2010) and Discrete Cosine Transform (DCT-EnKF) (Jafarpour et al. 2008) are some of these methodologies in which the model parameters are transformed into other Gaussian variables, which can be assimilated within the EnKF framework.

Aside from matching the dynamic flow observations, in order to successfully characterize a complex reservoir, it is important to condition the predicted models to reference geologic information such as conceptual models, log, core, seismic interpretations, and statistics from similar fields. Such data are generally available for reservoir modeling and should be incorporated in the studies. Therefore, characterization of a robust description of geological features such as facies distributions (proportions and spatial patterns) and channel properties (e.g. orientation, widths, and amplitudes) should incorporate all the static information and honor the reference statistics (e.g. histogram, spatial pattern and variogram or covariance) inferred from the reference geologic information.

While predicting model parameters using EnKF, the initial ensemble is designed to account for all uncertainties in the reservoir model. A major portion of the ensemble uncertainty will be diminished after a few EnKF update steps (Jafarpour et al. 2011), and realizations with poor initial estimates undertake major changes to match early production history. As a result of this early update, the spatial relation of the properties is destructed, and the model parameter variogram will not follow the geological properties of the formation, on which the initial ensemble was generated. The calculated experimental variogram along different azimuths would deviate from the reference benchmark variogram (inferred from static geologic information), and the short-scale variability was systematically larger for all the updated realizations.

The motivation of this chapter is to present a modified approach to honor the reference statistics (i.e. histogram, spatial pattern and variogram) within the

conventional framework of EnKF using DCT parameterization and by incorporating a re-sampling step. After the initial EnKF assimilation steps, the updated ensemble members are more consistent with the early production history. A re-sampling procedure is implemented to generate an "improved" ensemble of model parameters by P-Field (Srivastava 1992) simulation. The probability distributions at all grid block locations and the spatial correlation of the probability fields are used in P-Field simulation. The new ensemble incorporates the updated model information from the initial steps and consists of realizations that are more consistent with both the dynamic and static data.

2.2 Methodology

In this section, various components of the implementation are discussed in detail. Specific modifications to the conventional EnKF procedure are highlighted.

2.2.1 Generation of an initial ensemble of facies indicators

The conditional Sequential Indicator Simulation (SIS) (Deutsch and Journel 1997) is implemented to generate an initial ensemble of facies distribution. Facies observations at the well locations are used as the conditioning data, and geostatistical parameters (histogram and variogram model) are typically inferred from other regional geologic information

$$F(z) = E\{I(u;z)|(n)\} = P\{Z(u) \leq z|(n)\}, \forall u \in Domain$$

$$I(u;z) = \begin{cases} 0 & if \quad Z(u) \leq z \\ 1 & if \quad Z(u) > z \end{cases}$$

<div align="right">2-1</div>

Consider the random function $Z(u)$, $u \in R^n$ in a discretized grid system with N grids. A set of conditioning data $d_h = \{d(u_j), j = 1, ..., h\}$ represents the hard data available from well information. The set of data for conditional simulation including the previously simulated nodes is represented by Λ_i for every location u_j and $\Lambda_0 = \{d_h\}$. The conditional simulation in the grid system with N grids is based on sampling from the N-variate distribution conditioned on the set Λ_0

$$F(u_1, ..., u_N; z_1, ..., z_N | \Lambda_0) = P\{Z(u_1) \leq z_1, ..., Z(u_N) \leq z_N | \Lambda_0\}$$ 2-2

In this study, it has been assumed that the domain is stationary. Facies proportions are assumed to be known, and the same proportions are used to generate all ensemble members of the case studies. Furthermore, the spatial pattern and continuity of the facies are defined by two-point covariance model (or variogram) and multiple-point statistics extracted from a training image in the case studies.

2.2.2 Forecast model

The forecast model is separately applied to each ensemble member. The multiphase fluid flow in petroleum reservoirs is described by material balance, momentum balance, phase behaviour descriptions and numerous auxiliary equations. This system of non-linear differential equations is solved numerically using methods such as finite difference

$$\begin{bmatrix} u_k^p \\ d_k^p \end{bmatrix} = g \begin{bmatrix} m_{k-1}^a \\ u_{k-1}^a \end{bmatrix},$$
$$\text{2-3}$$

In the above equation, m and u denote the model and state variables, respectively, and d_k represents the predicted production data at k^{th} step. The superscripts, p and a, specify the predicted and analyzed states, respectively. In this study the commercial reservoir simulation software Eclipse black-oil (2011) was used as the forecast model.

2.2.3 DCT parameterization of non-Gaussian facies indicators for continuous reservoir model updating in EnKF (DCT-EnKF)

The Discrete Cosine Transform (DCT) has its roots in image processing (Jain 1988; Rao and Yip 1990). It is widely used in signal and image processing for data compression. The one-dimensional DCT (X_k) of a signal x_n of length N is defined as

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}(n+1/2)k\right] \qquad k = 0,...,N-1 \qquad \text{. 2-4}$$

The inverse DCT is defined as

$$x_n = \sum_{k=0}^{N-1} X_k \cos\left[\frac{\pi}{N}(n+1/2)k\right] \qquad n = 0,...,N-1 \qquad \text{. 2-5}$$

In our studies, DCT is implemented to parameterize the facies indicators into coefficients of the retained cosine basis functions. These coefficients are incorporated into the state vector and updated in the EnKF procedure. At the

end, they are transformed back into facies indicator values by means of inverse DCT transform (Jafarpour et al. 2008). In order to reduce the size of the model parameter space and capture the spatial continuity of different facies, few DCT coefficients are included in the state vectors. For practical applications, other parameterization techniques, which were mentioned earlier, can be applied instead of DCT. If the history-matched models contain artificial short-scale variability, smoothing algorithms should be implemented (Nejadi et al. 2012a, b).

2.2.4 Implementation of the analysis step

The analysis step is carried out using the observation data (d_k) for all ensemble members

$$y_k^{a,j} = y_k^{p,j} + K_k \left(d_{ik} - H y_k^{p,j} \right),$$ 2-6

where $y_k^{a,j}$ represents the analyzed state vector, $y_k^{p,j}$ is the posterior state vector, K_k is the Kalman Gain at k^{th} step, and H is the observation operator, which represents the relationship between the state vector and the observation vector:

$$H = \begin{bmatrix} 0 \mid I \end{bmatrix},$$ 2-7

d_{ik} is the observation data at the kth step ($d_{obs,k}$) plus observation noises (ε_k) for ensemble member i

$$d_{ik} = d_{obs,k} + \varepsilon_k,$$ 2-8

Kalman gain K_k is defined as

41

$$K_k = C_{y_k^{p,j}} H^T \left(H C_{y_k^{p,j}} H^T + C_{d_k} \right)^{-1},$$
<div align="right">2-9</div>

where $C_{y_k^{p,j}}$ is the state cross covariance matrix, and C_{d_k} is the error covariance matrix.

The cross covariance matrix is approximated as

$$C_{y_k^{p,j}} \approx \frac{1}{N_e - 1} \left\{ \left(y_k^p - \overline{y_k^p} \right) \left(y_k^p - \overline{y_k^p} \right)^T \right\},$$
<div align="right">2-10</div>

in which T indicates matrix transpose, and $\overline{y_k^p}$ is the average of all posterior state vectors.

The forecast and update steps are repeated as new observations become available.

2.2.5 Re-sampling of the ensemble with probability maps

In this section, a novel re-sampling procedure is proposed to address two common issues associated with conventional EnKF implementation: (1) honoring of reference statistics of the model parameters during model updates and (2) improving ensemble variability. In this approach, after certain number of EnKF assimilation steps, an "improved" ensemble is generated. In order to maintain the diversity among ensemble members, a re-sampling procedure is needed to generate a new ensemble stochastically from assimilated members. In this new approach, a probability distribution map is derived from the assimilated ensemble and is further implemented in probability field simulation. Performing

probability field sequential simulation ensures that model parameter statistics and the corresponding spatial relationship are captured. Following the re-sampling procedure allows that the new ensemble incorporates the updated model information from the initial EnKF assimilation steps and that the re-generated ensemble is consistent with the static geological data. In order to warrant that the rest of the model state states (i.e., fluid saturations and pressure) are consistent with the hydraulic conductivity of the re-sampled ensemble, the new members are then subjected to the forward model (reservoir simulator) from the beginning (initial conditions) to the last update step prior to re-sampling. After that, updating would resume for the remaining production history using EnKF.

The probability or P-Field simulation introduced by Srivastava (1992) has gained popularity because of its simplicity, its capability in accounting for secondary information and its computational efficiency (Froidevaux 1993; Pyrcz and Deutsch 2001). P-Field simulation involves two separate steps. Firstly, local probability density function is defined for all grid block location to be simulated. The probability distributions are a step function at the well locations or at the locations where hard data is available. In this study, facies indicators (categorical variables) are modeled, and the local probabilities (pdf) of model parameters are constructed from the assimilated ensemble after certain steps. In the next step of P-Field simulation algorithm, a set of spatially correlated probability values is generated by Monte Carlo simulation (in the interval [0, 1]) and used to sample simulated values from those local probability distributions. The probability

values have a uniform distribution and should not honor any conditioning hard data.

A drawback of the P-Field simulation is that the simulated models have greater continuity than expected, resulting in less accurate production performance than expected. However, further data assimilation using EnKF would remedy this excessive continuity. An alternative methodology would be to model the probability distributions as a trend. In that case, both the trend and the residual should be modeled; however trend modeling would not reproduce the expected spatial continuity for some specific complex reservoir models such as channelized formations.

2.2.6 Choice of re-sampling time

The proper selection of the number of assimilation step(s) or time at which re-sampling is performed is discussed here. As the number of EnKF assimilation steps increases, the root mean square error (RMSE) of the predicted data mismatch (objective function) decreases. A new ensemble is re-sampled from assimilated ensemble members that are updated in accordance with the early production data; therefore, the probability map used for re-sampling should be calculated after the objective function has experienced a moderate decline. The objective function is defined as the dynamic data mismatch.

$$O(m) = [g(m) - d_{obs}]^T C_D^{-1} [g(m) - d_{obs}] \approx RMSE ,$$

<div align="right">2-11</div>

where the simulation results (data predictions) $g(m)$ are compared with observed data (d_{obs}), and C_D is the error covariance matrix.

The ensemble variance would diminish after several assimilation steps. In order to avoid ensemble collapse and maintain diversity among the re-sampled ensemble members, this re-sampling should be performed while the ensemble variance is still high. At the beginning of the assimilation procedure, the mismatch is high, reflecting the uncertainty in the initial ensemble and the respective proximity of each member to the true dynamic responses. As the assimilation procedure continues, the objective function $O(m)$ (or RMSE) decreases. A major portion of the ensemble uncertainty is corrected after the first few assimilation steps (Jafarpour et al. 2011).

After several update steps, model parameter distribution of the members with poor initial estimations cannot be further improved to reduce the objective function $O(m)$. In contrast, the ensemble member diversity incessantly decreases after each and every update step, further reducing the ensemble variance. Re-sampling carried out at that point would yield an underestimation of the posterior pdf of the model parameters (Emerick and Reynolds 2012). As a result, the proper re-sampling step is where the objective function has reasonably decreased and, at the same time, the ensemble members reflect sensible diversity.

2.3 Implementation of P-Field Simulation

The proposed methodology has been implemented for the history matching of two synthetic reservoir models: one with three facies and another channelized model with two facies. The objective is to characterize the facies distributions by conditioning to both observed production data and static measurements at the wells, in addition to maintaining the spatial correlation of model parameters and reproducing the reference histogram and spatial continuity of model parameters.

2.3.1 Model description – Example A

The numerical model for this case study consists of a 21 × 21 grid in the x-y direction. The dimensions of the model are 1600 m × 1600 m × 6 m. Nine vertical wells are located in the reservoir with one oil producer, four water injectors, and four pressure observations wells (Figure 2-1).

Three distinct facies, namely sand (S), fine sand (FS) and shale, are present in the model. Facies distribution for the reference model is generated using unconditional sequential indicator simulation (SIS), as implemented in GSLib (Deutsch and Journel 1997). The reference facies proportions are 45% S, 37% FS and 18% shale. Based on the geological properties of the reservoir obtained from regional studies, an experimental variogram is inferred as an exponential variogram model with maximum and minimum correlation lengths of 1370 m and 518 m, respectively, along an azimuth angle of 135°. Permeability within each facies was assumed to be uniform with an average value (2000 md, 500 md, and 50 md for S, FS, and shale, respectively), and rock porosity of different

facies were assumed to be constant (0.3) throughout the model. Areal facies distribution of the reference model is depicted in Figure 2-2a.

The production schedule is defined for 32 years. The first 10 years of production data are used for history matching, and the complete production history (32 years) is used to evaluate the predictions and calculate the RMSE. The production and injection wells are operated at constant bottom-hole pressure. The synthetic field-observed data set, which consists of oil production rate, gas oil ratio, and water cut of the producer, water injection rates, and bottom-hole pressure of the observation wells, is obtained by subjecting the reference model to a reservoir simulator. Table 2-1 presents the important model parameters of the reference model.

2.3.2 Base case – Example A

This case is intended to examine the performance of the DCT-EnKF method for assimilating model parameters of a multiple facies model. Initially, 100 ensemble members are generated using conditional SIS. The map of average permeability of the initial ensemble is depicted in Figure 2-3a. The facies observations at all 9 well locations are used as the conditioning data. The variogram parameters and the reference distribution are the same as the benchmark case. Facies proportions of the simulated realizations were within a range of 5% of the benchmark case.

History matching was performed for 20 steps. Assimilation steps are every six months. The map of average permeability of the updated ensemble (after 20 assimilation steps) is shown in Figure 2-3b. Permeability map of a randomly

47

selected ensemble member is also shown in Figure 2-4, before and after assimilation. It is apparent that despite the production history being matched, the spatial continuity of facies distribution in the updated ensemble deviates significantly from the benchmark case. The production and injection profiles before and after history matching are shown in Figure 2-5, while summaries the RMSE of the production match on a 0 to 1 scale are given in Table 2-2.

2.3.3 Probability re-sampling – Example A

In this case study, the DCT-EnKF updated facies realizations obtained after certain assimilation steps of the base case are used in the P-Field simulation to generate a new ensemble of facies realizations. In particular, after three assimilation steps (18 months), the ensemble members are used to calculate the pdf of facies indicators, which are subsequently used to generate new ensemble members in P-Field simulation.

The re-sampled ensemble is subjected to the forward model (reservoir simulator) from the beginning to the last update step prior to re-sampling (18 months). The DCT-EnKF is applied for the remaining production history (17 assimilation steps, 8.5 years of production history). The map of average permeability (before and after the update) shows significant improvement (Figure 2-6), as compared to the base case (Figure 2-3), in terms of exhibiting similar spatial characteristics as depicted by the reference distribution and the benchmark case. The permeability map of a randomly selected ensemble member is shown in Figure 2-7 for before and after the update. The production and injection profiles

before and after history matching are compared in Figure 2-8. The corresponding RMSE values of the history-match results are summarized in Table 2-2.

The semi-variograms of the base case, P-Field re-sampling, and the true model along minor and major directions of continuity (the horizontal azimuths of 45° (a) and 135° (b)) are presented in Figure 2-9. Facies histograms for different cases are also compared in Figure 2-10.

2.3.4 Choice of re-sampling time – Example A

Results of the objective function (from Equation 2-11) in terms of production data match at different assimilation steps are shown in Figure 2-11. The value is high at the beginning, illustrating the uncertainty in the initial ensemble. After a few assimilation steps, the value decreases but remains relatively constant until water breakthrough. As discussed previously, a major portion of the ensemble uncertainty has been corrected after the few initial assimilation steps, which is in agreement with observations made by Jafarpour et al. (2011). Also shown in Figure 2-11 is the normalized ensemble variance of model parameters at a few selected assimilation steps. In between the third and thirteenth step, the value is constant (Figure 2-11), implying that EnKF assimilation has not improved the quality of the history match; however, the ensemble variance continues to diminish. Given that the proper re-sampling step is where the objective function (RMSE) has decreased substantially without compromising the ensemble variability, it is concluded that for this example, re-sampling should be carried out after the third assimilation step.

2.3.5 Model description – Example B

The numerical model for this case study consists of a 100 × 80 grid in the x-y direction. The dimensions of the model are 1000 m × 800 m × 10 m. Two horizontal line injectors and producers equipped with inflow control valves are located in the reservoir. Well locations are shown in Figure 2-1.

Two distinct facies, namely sand (S) and shale, are present in the model. Facies distribution for the reference model is adapted from the channelized model presented in Li et al. (2012). The reference facies proportions were 30% S and 70% shale. Permeability within each facies was assumed to be uniform with an average value (1000 md and 50 md for sand and shale, respectively), and porosityis assumed to be constant (0.2) throughout the model. Spatial facies distribution of the reference model is depicted in Figure 2-2b.

The production schedule is defined for 16 months. The production and injection wells are operated at constant bottom-hole pressure. A synthetic production data set, which consists of oil production rate and water injection rates, is obtained by subjecting the reference model to reservoir simulator. The Eclipse black-oil simulator is used as the reservoir simulator (forward model). The important model parameters of the reference model are presented in Table 2-1.

2.3.6 Base case – Example B

The DCT-EnKF method is applied for data assimilation of this two-facies channelized reservoir model. Initially, 100 ensemble members are generated by multiple-point Single Normal Equation Simulation (SNESim), as implemented

in SGeMS (Remy et al. 2009). Facies classifications at the horizontal well locations are used as the conditioning data, and Figure 2-12 shows the training image implemented in SNESim. The map of average permeability of the initial ensemble is depicted in Figure 2-13a.

Five months of production data are used for history matching with an assimilation step of one month. The map of average permeability of the updated ensemble (after five assimilation steps) is shown in Figure 2-13b. Permeability map of a randomly selected ensemble member before and after update is shown in Figure 2-14. The RMSEs of the production match are summarized in Table 2-2. It is apparent that despite the production history being matched, the spatial continuity of facies distribution in the updated realizations is different from the reference model and the training image.

2.3.7 Probability re-sampling – Example B

The re-sampling scheme is applied after the fifth assimilation step. Same as Example A, the ensemble members are used to calculate the pdf of the facies indicators. The approximated local pdf is used subsequently to generate a new ensemble using SNESim. The new ensemble (100 members in total) is subjected to the forward modeling (for five months) and is further assimilated within the DCT-EnKF framework to match the remaining 11 months of production history. The RMSE values of the history-match results are summarized in Table 2-2. Average permeability distributions of the initial and final updated ensemble are shown in Figure 2-15. The updated ensemble is acceptable in terms of RMSE values, and the resulting permeability map has shown significant improvement,

as compared to the base case (Figure 2-13), in terms of exhibiting similar spatial characteristics as depicted in the training image (Figure 2-12) and the true case (Figure 2-2b). Permeability map of a randomly selected ensemble member is shown in Figure 2-16 for before and after the update.

2.3.8 Choice of re-sampling time – Example B

The calculated objective function (from Equation 2-11) in terms of production data match at different assimilation steps is presented in Figure 2-17. In a fashion similar to the previous case study – Example A, its value is high at the beginning due to uncertainty in the initial ensemble). After a few assimilation steps, the value decreases to a stabilized value and increases again at water breakthrough. Also shown in Figure 2-17 is the normalized ensemble variance of model parameters at a few selected assimilation steps. This value remains primarily constant between the fifth and eighth step (Figure 2-17), implying that EnKF assimilation has not improved the quality of the history match; however, the variability among ensemble members continues to diminish. It is concluded that, for this example, re-sampling should be carried out after the fifth assimilation step.

2.4 Results and Discussion

Different parameters are considered and taken into account in our studies in order to evaluate the performance of the methodology. First, the RMSE values are used in all case studies to measure the quality of the observed data match. Second, in Example A, experimental variogram of the benchmark and the

updated realizations are calculated along the minor and major directions of continuity. The calculated values for the small lag distances and the variogram range in minor and major directions of continuity are compared. It is noteworthy that all variogram data are standardized to a sill of one for our studies. For the second model (Example B), the continuity of the sand facies and reproduction of the correct spatial pattern (same as the training image shown in Figure 2-12) are taken into account. Finally, the facies proportions of the updated realizations and the true model are compared for different case studies to verify that the correct proportions have been maintained in the posterior ensemble.

The RMSE values are calculated for all cases, and the averages of the calculated RMSE values for all realizations are presented in Table 2-2. The results clearly demonstrate that an acceptable production history match is achieved for the re-sampling case, and the results are greatly improved as compared to the base cases.

In each of the examples, it is illustrated through the comparison with the base case that, although production history matching was achieved after several assimilation steps, the spatial continuity of geologic features were no longer preserved with conventional EnKF. This is evident as we compare Figure 2-3b and Figure 2-4b with Figure 2-6b and Figure 2-7b. It is observed that spatial continuity in the model parameters is more accurately reproduced in the updated ensemble of the P-Field case when compared against the base case. The same conclusion can be deduced in the channelized example, where facies continuity is destroyed after several update steps (Figure 2-13 and Figure 2-14).

53

Figure 2-15b illustrates how the re-sampling scheme has successfully re-established this spatial continuity in the final updated ensemble. Nonetheless, the obtained results clearly demonstrate the importance and performance of the proposed re-sampling technique and its potential in modeling the non-Gaussian geological features in reservoirs.

Prior to any dynamic updating, few hard data are typically available to generate the initial ensemble. Many researchers have reported the issue of ensemble collapse due to a poor initial ensemble. The ensemble typically experiences major updates during the first few steps and immediately after a new type of data becomes available, as illustrated in Figure 2-11 and Figure 2-17, which show the objective function (Equation 2-11) value decreases with successive update steps up to the thirteenth step in Example A and ninth step in Example B. The varying breakthrough characteristics among different ensemble members cause the objective function to increase sharply. Thereafter the model parameters are updated in accordance to the water cut measurements, and the objective function starts to decline once again.

2.5 Conclusion

A novel re-sampling procedure was proposed to (1) correct for the loss of non-Gaussian contributions in model parameters after the ensemble Kalman filter update and to (2) honor the reference distribution obtained from static geologic information. The method entails combining probability field re-sampling with the discrete cosine transform - ensemble Kalman filter approach. The results have shown the success of the re-sampling technique in model characterization

54

of multiple (up to three) facies. The observed dynamic data were matched, and the geological feature of the reservoir in terms of reference facies proportions and spatial distribution of model parameters (variogram, training image) were honored. In order to ensure consistency, after every re-sampling, the forward model (reservoir simulation model) should be applied from the beginning (zero time step). This would certainly incur additional computational efforts. However, the proposed re-sampling procedure presents a promising potential to model non-Gaussian model parameters. It is also important to note that the proposed re-sampling is carried out after the first few assimilation steps, in contrast to most iterative ensemble Kalman filter formulations proposed by other researchers in the past.

Table 2-1 Synthetic reservoir model parameters.

Example A	
Model Dimensions	$21 \times 21 \times 1$ grids in the X, Y and Z directions
Grid Dimensions	76 m × 76 m × 6 m (250 ft × 250 ft × 20 ft)
Reservoir Depth	914 m (3000 ft)
Porosity	30% constant
Permeability Range	50 md, 500 md and 2000 md
Operating Bottom-hole Pressure	Producer: 15.8579 MPa (2300 psi) Injectors: 17.2369 MPa (2500 psi)
Noise Standard Deviation	Producer: 7.94937 m3 (50 STBPD[2]) Injectors: 3.17975 m3 (20 STBPD)
Initial Reservoir Pressure	17.2369 MPa (2500 psi)
Total Generated History	32 years (10 years was used for history matching)
Example B	

---

[2] Stock Tank Barrel Per Day

Model Dimensions	100 × 80 × 1 grids in the X, Y and Z directions
Grid Dimensions	10 m × 10 m × 10 m (33 ft × 33 ft × 33 ft)
Reservoir Depth	914 m (3000 ft)
Porosity	20% constant
Permeability Range	50 md and 1000 md
Operating Conditions	
Producers	BHP[3]= 19.3053 MPa (2800 psi)
Injectors	BHP = 22.0632 MPa (3200 psi)
Noise Standard Deviation	Producers: 3.17975 m3 (20 STBPD) Injectors: 3.17975 m3 (20 STBPD)
Initial Reservoir Pressure	20.6843 MPa (3000 psi)
Total Generated History	16 months

Relative Permeability and Capillary Pressure Curves			
Sw	Krw	Kro	Pcow
0.2	0	1	7
0.3	0.07	0.4	4
0.4	0.15	0.125	3
0.5	0.24	0.649	2.5
0.6	0.33	0.0048	2
0.8	1	0	0

Table 2-2 Root mean square values (on a 0 to 1 scale, where 1 represents a perfect match) of observations after history matching for different cases.

Example-A	Producer (Oil production rate)	Injector (Water injection rate)				Pressure Observation (BHP)			
Well No.	1	1	2	3	4	1	2	3	4
Base Case	0.991	0.788	0.983	0.938	0.972	0.874	0.978	0.952	0.928
P-Field	0.995	0.997	0.872	0.965	0.982	0.943	0.977	0.971	0.951
Example-B									
Wells	Producers (Oil production rate)					Injectors (Water injection rate)			
Base Case	0.867					0.749			
P-Field	0.971					0.987			

---

[3] Well Bottom Hole Pressure
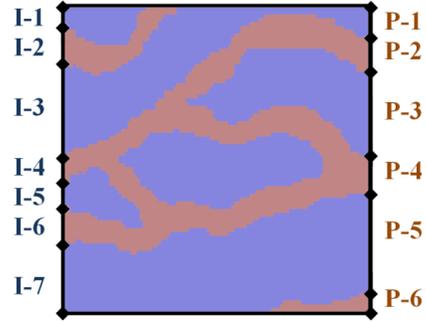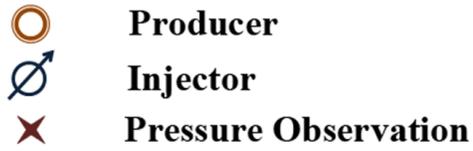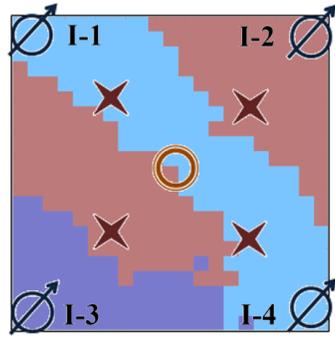
I-1 to I-7 : Horizontal Injectors

P-1 to P-6 :Horizontal Producers

Figure 2-1 Well locations.
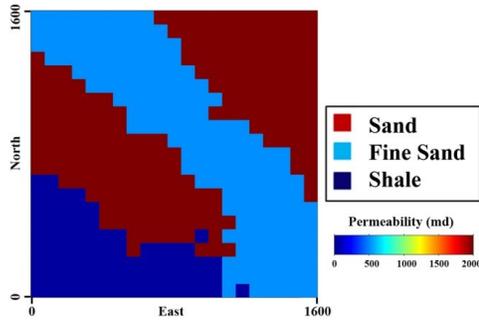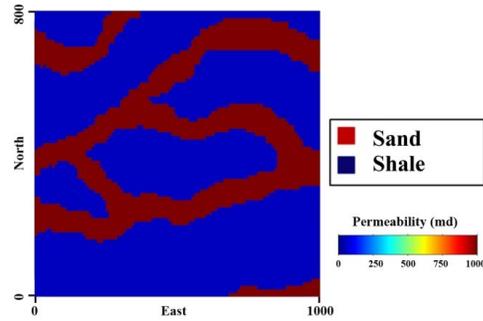
(a)   (b)

Figure 2-2 Facies distribution of the true model for: (a) Example A and (b) Example B. The units for X-axis (East) and Y-Axis (North) are in meters.

Base Case - Example A
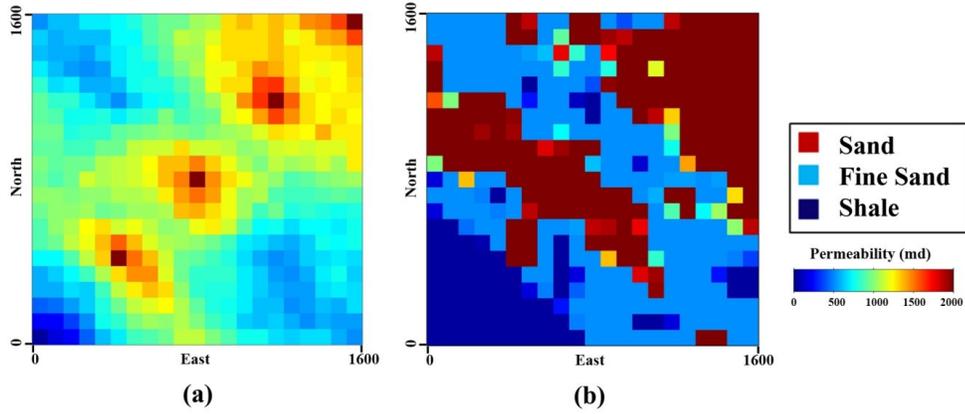
Figure 2-3 Average permeability map of the (a) initial and (b) updated realizations (20 update steps, 10 years of production history). The initial realizations were generated using conditional Sequential Indicator Simulation. The units for X-axis (East) and Y-axis (North) are in meters.

Base Case - Example A
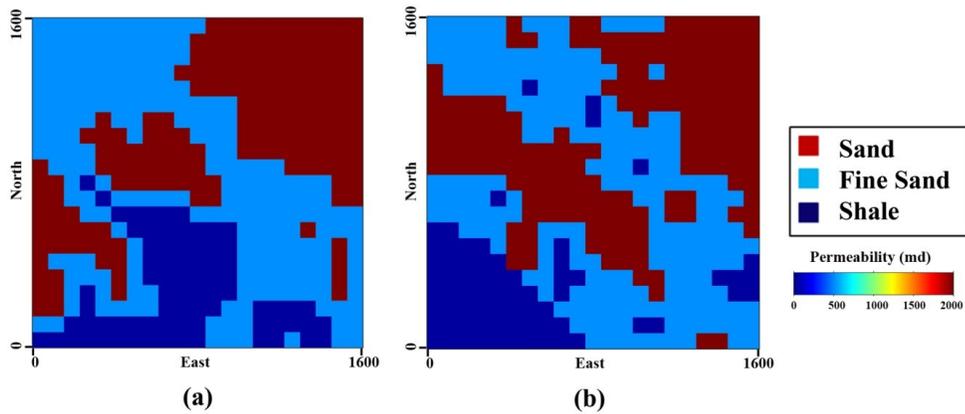
Figure 2-4 Permeability map of the (a) initial and (b) updated model for a single realization. The initial realization was generated using conditional Sequential Indicator Simulation. The units for X-axis (East) and Y-Axis (North) are in meters.

58
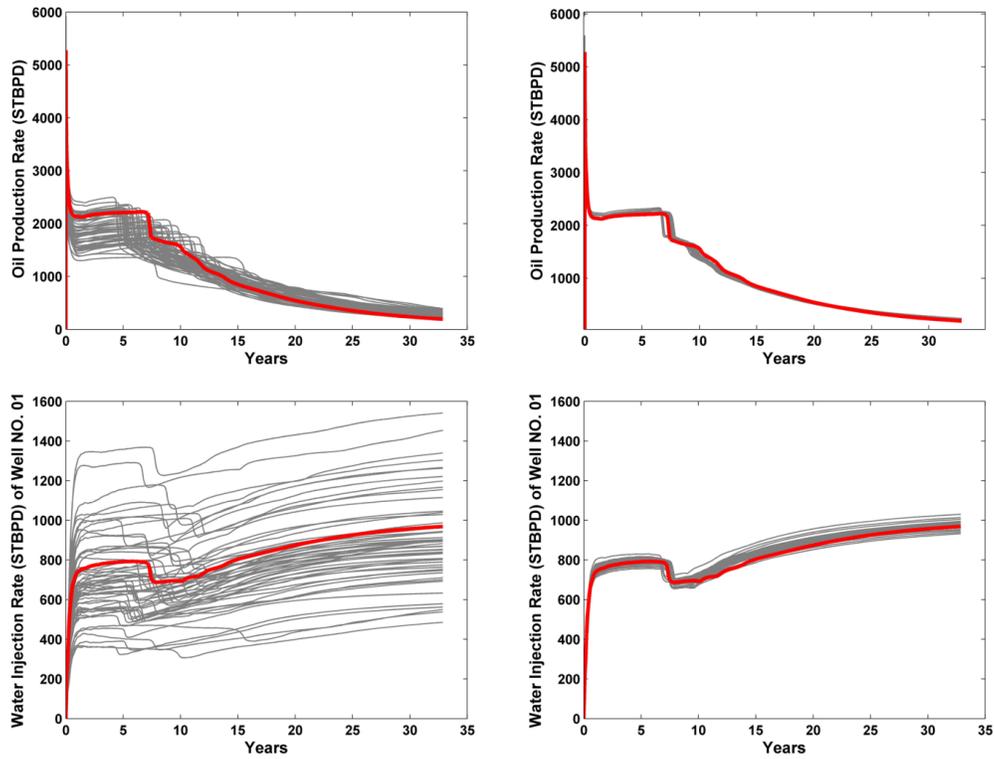
Figure 2-5 Oil production rate and water injection rates in stock tank barrel per day (STBPD) before (left column) and after (right column) history matching for the base case, Example-A. The red line represents the production history, and the grey lines show the behavior of different realizations.

Case PField - Example A

(a)   **(b)**
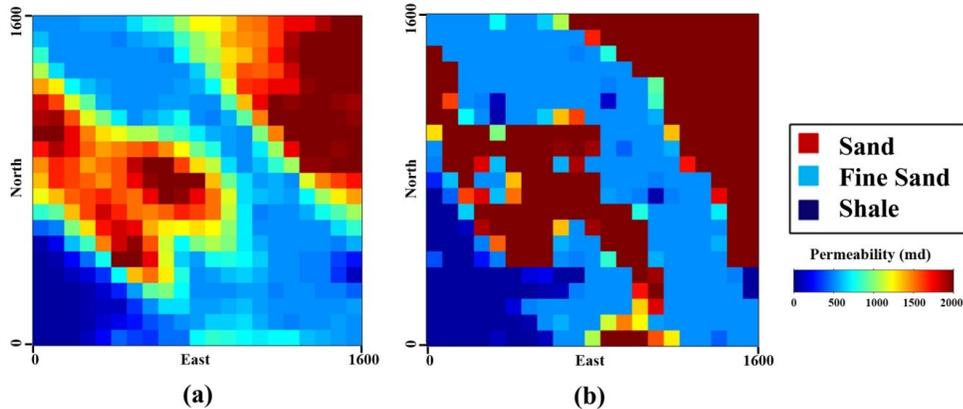
Figure 2-6 Average permeability map of the (a) initial and (b) updated realizations (17 update steps, 8.5 years of remaining production history). The initial ensemble was generated using the P-Field re-sampling method (based on the results of the base case after

59

three update steps, 1.5 years of production history). The units for X-axis (East) and Y-Axis (North) are in meters.

Case PField - Example A

(a)   (b)
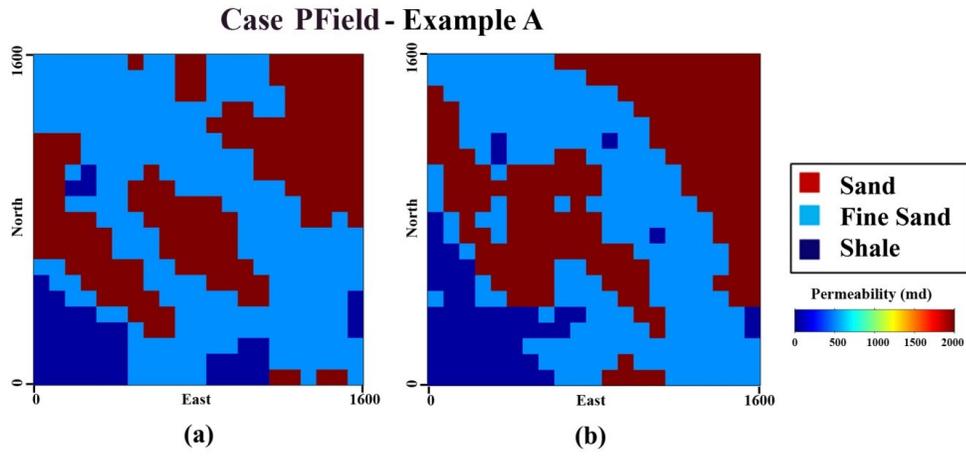
Figure 2-7 Permeability map of the (a) initial and (b) updated model for a single realization. The initial realization was generated using the P-Field re-sampling method. The units for X-axis (East) and Y-Axis (North) are in meters.
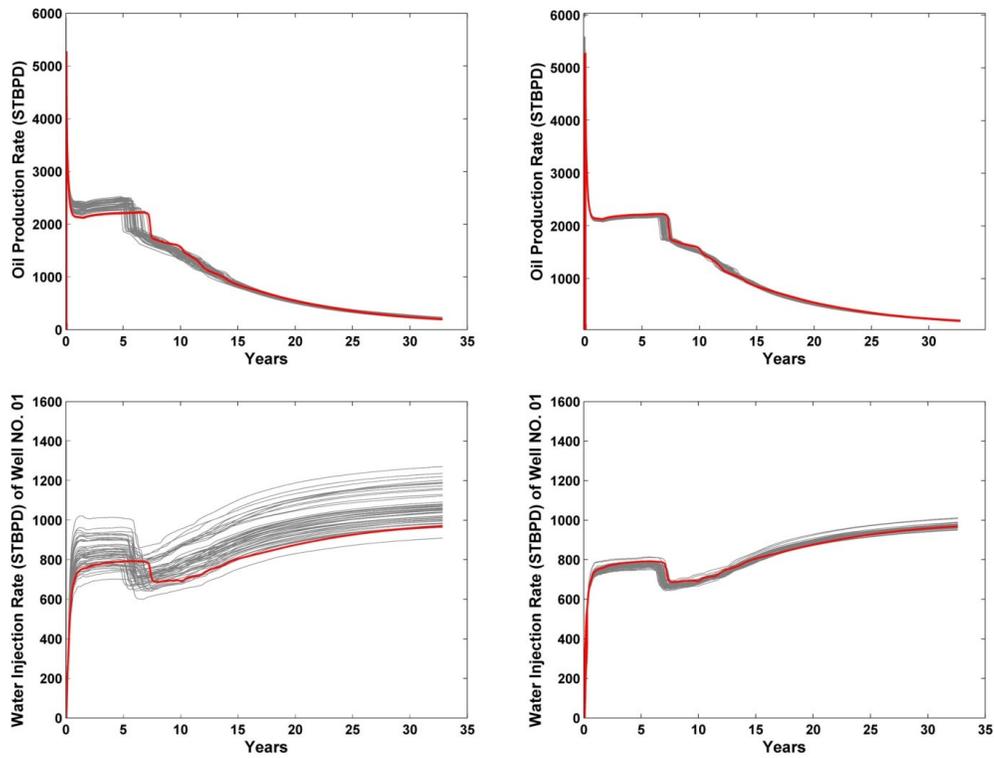
Figure 2-8 Oil production rate and water injection rates in stock tank barrel per day (STBPD) before (left column) and after (right column) history matching for P-Field re-sampling, Example-A. The red line represents the production history, and the grey lines show the behavior of different realizations. The initial ensemble was generated using the P-Field re-sampling method.
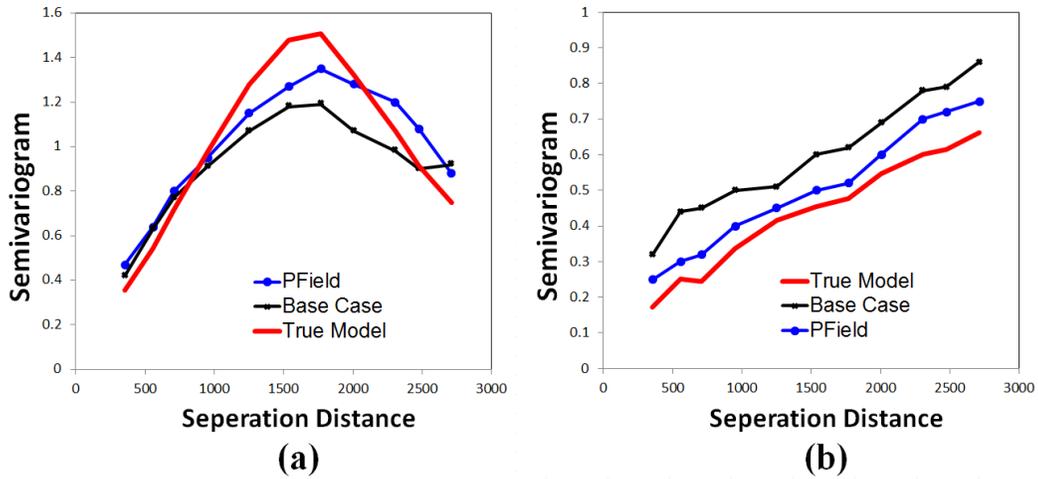
Figure 2-9 Experimental variograms of the top updated ensemble member (smallest RMSE) for the base case and re-sampling with P-Field compared with the true model, along the horizontal azimuths of (a) 45° and (b) 135° (counter clockwise from east). The unit for X-axis (Separation Distance) is in meters.
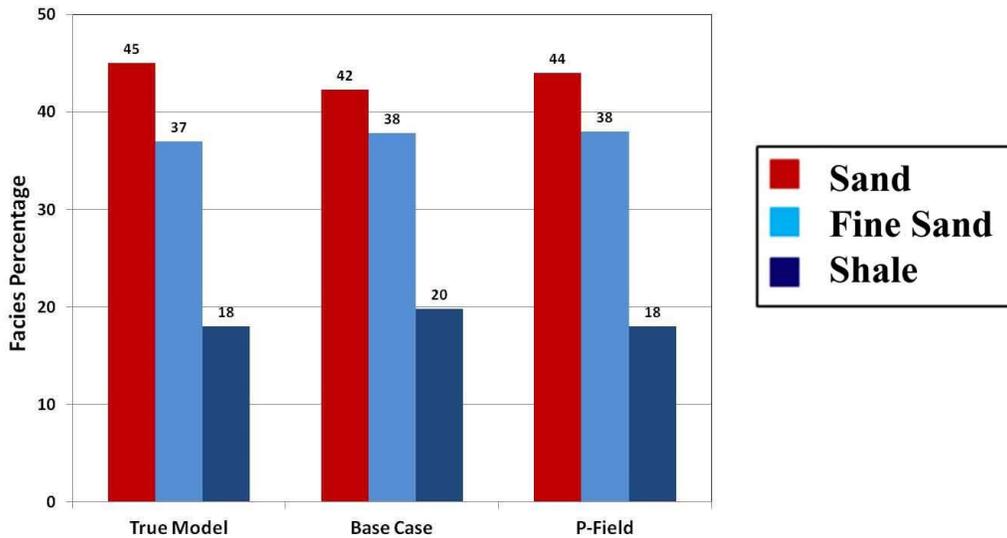
Figure 2-10 Facies histogram for different cases of Example A compared to the true model's facies proportions.
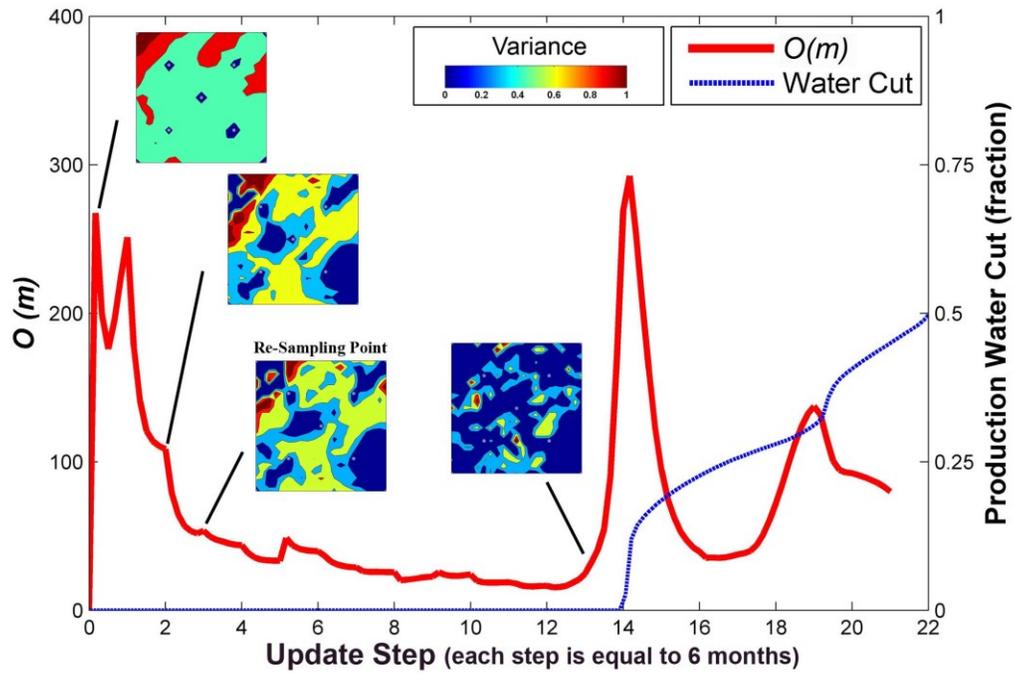
Figure 2-11 Objective function defined by Equation 2-11 (red line) and the true production water cut (water cut) at successive assimilation steps. The inset figures show the ensemble variances of model parameters at various assimilation steps: t = 0, 12, 18 and 78 months.
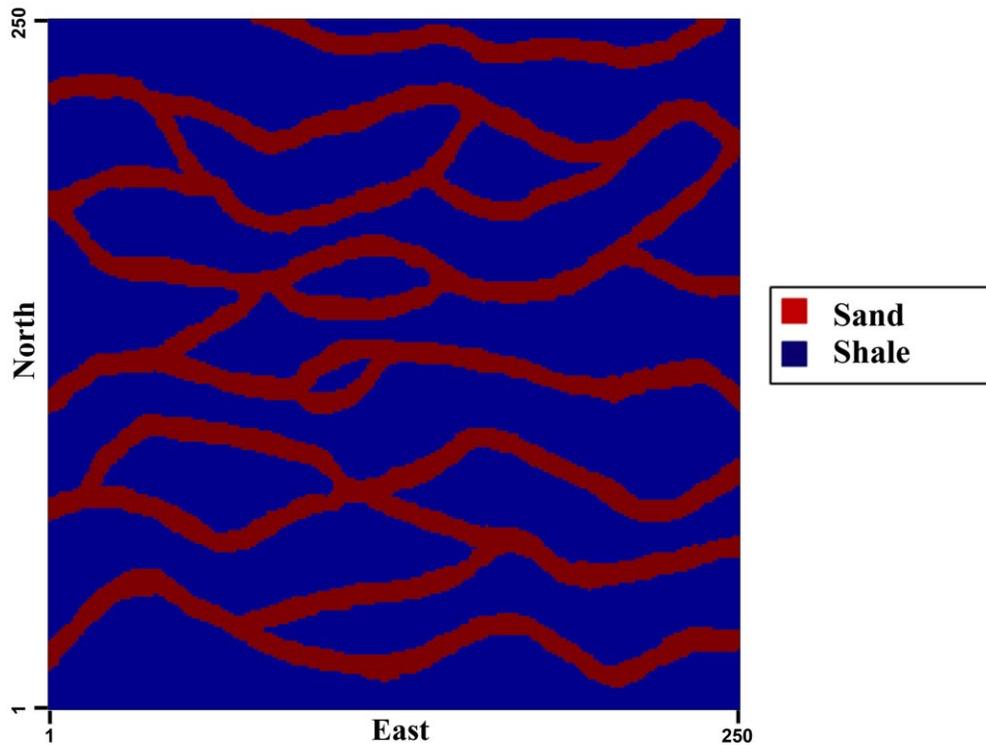
Figure 2-12 The training image implemented in Single Normal Equation Simulation. Dimension of the training was 250 by 250 grids.

Base Case - Example B

(a) (b)
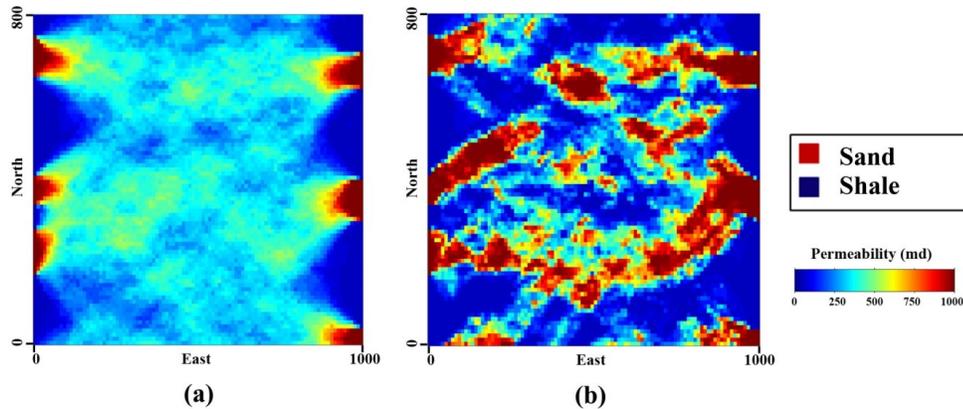
Figure 2-13 Average permeability map of the (a) initial and (b) updated realizations (five update steps, five months of production history). The initial realizations were generated using Single Normal Equation Simulation. The units for X-axis (East) and Y-Axis (North) are in meters.

63

Base Case - Example B

(a)   (b)
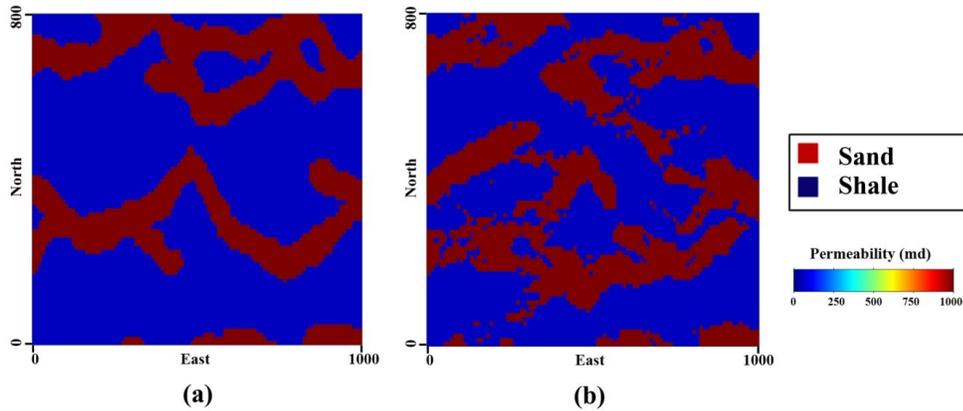
Figure 2-14 Permeability map of the (a) initial and (b) updated model. The initial realization was generated using Single Normal Equation Simulation. The units for X-axis (East) and Y-Axis (North) are in meters.

Case PField - Example B

(a)   (b)
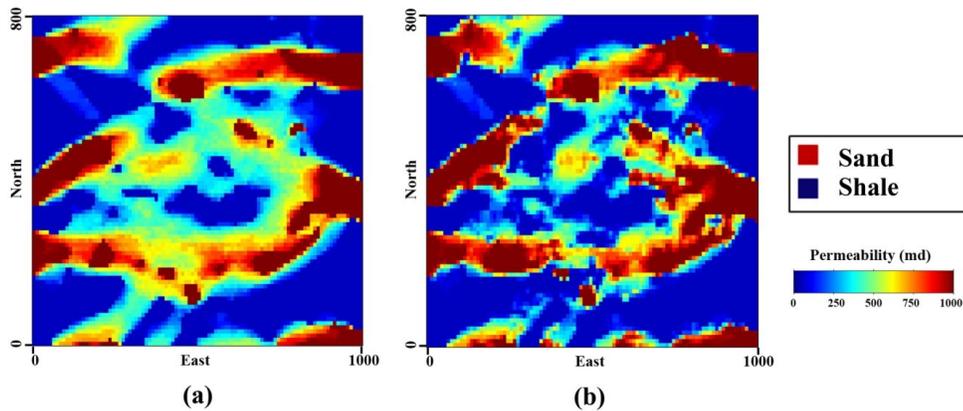
Figure 2-15 Average permeability map of the (a) initial and (b) updated realizations (11 update steps, 11 months of remaining production history). The initial ensemble was generated by the P-Field re-sampling method implemented in Single Normal Equation Simulation (based on the results of the base case after five update steps, five months of production history). The units for X-axis (East) and Y-Axis (North) are in meters.

64

Case PField - Example B
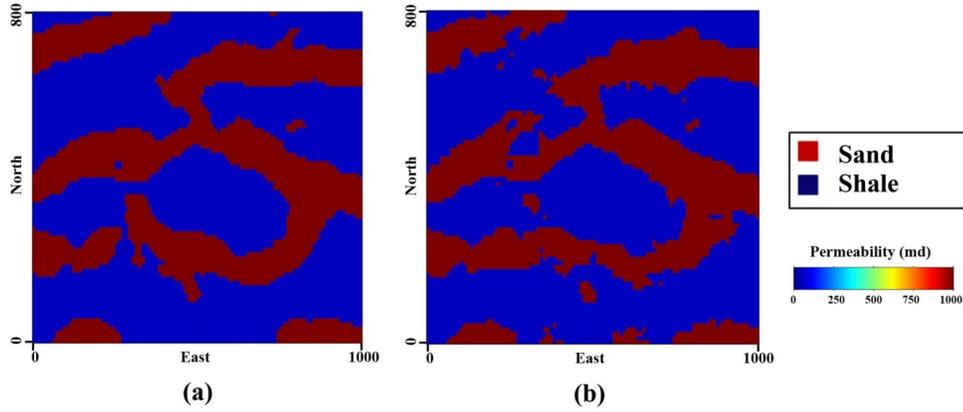
Figure 2-16 Permeability map of the (a) initial and (b) updated model. The initial realization was generated by the P-Field re-sampling method implemented in Single Normal Equation Simulation. The units for X-axis (East) and Y-Axis (North) are in meters.
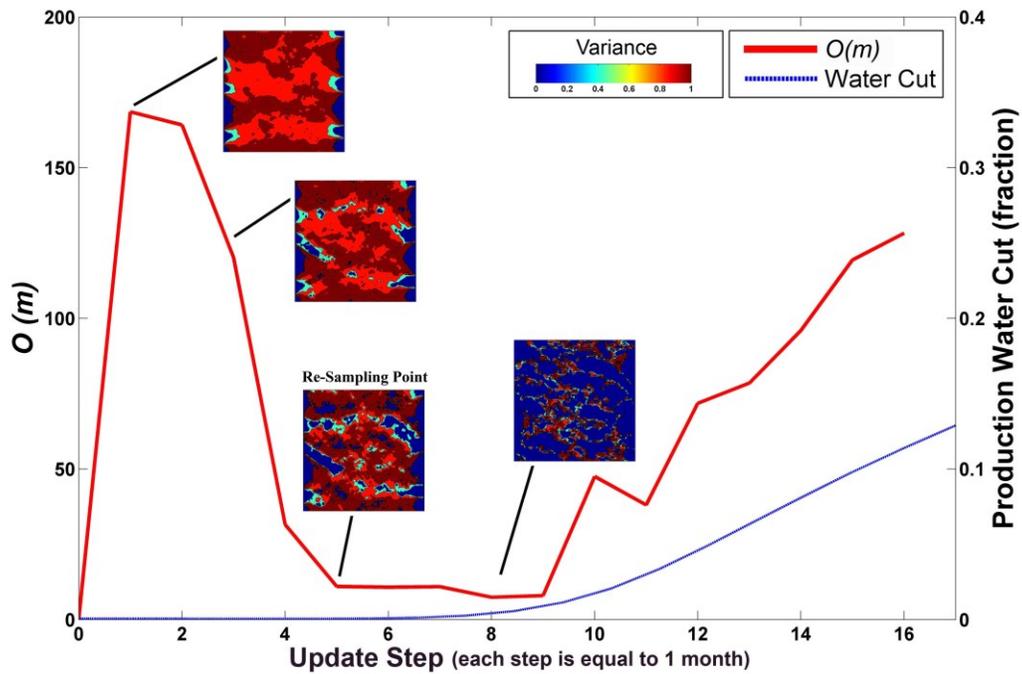
Figure 2-17 Objective function defined by Equation 2-11 (red line) and the true production water cut (blue line) at successive assimilation steps. The inset figures show the ensemble variances of model parameters at assimilation steps: t = 1, 3, 5 and 8 months.

Chapter 3: Characterization of Non-Gaussian Geologic Facies Distribution Using Ensemble Kalman Filter with Probability Re-Sampling[4]

3.1 Introduction and Background

Geostatistical methods have been commonly used to construct reservoir models by integrating data from diverse sources (Deutsch and Journel 1997). Many traditional methods rely primarily on static data (e.g., seismic, geological, well-log, and core data), which are sparsely located. Dynamic data (e.g., production rates, saturation, and pressure data), on the other hand, are abundantly available at well locations and could provide valuable information regarding the model parameters. The challenge of integrating such information, though, is that dynamic data is typically non-linearly related to the model parameters. The process of integrating all available data (both dynamic and static information) is referred to as history matching, with the objectives of quantifying heterogeneities and uncertainties in reservoir models, which would be used to obtain reliable production forecasts.

History matching of continuous rock properties such as porosity and permeability distributions in heterogeneous formations can be achieved using various data assimilation techniques. Many authors such as Oliver and Chen (2011) have presented thorough reviews of different available techniques in

---

recent years. However, as suggested by Saleri et al. (1992), for reservoirs exhibiting distinct facies distributions, a better approach is to parameterize the facies indicators (or rock types) directly as model parameters. The authors illustrated that given each facies represents a geologic unit with distinct flow behavior, fluid flow response is much more sensitive to perturbations in spatial arrangement and global proportions of discrete facies indicators than adjustment in spatial distribution of continuous model parameters such as porosity and permeability.

Numerous ensemble-based history matching techniques can be applied for updating realizations of facies indicators (Emerick and Reynolds 2013). These methods adopt an updating framework based on Bayes' theorem, where conditional (posterior) probability density function (pdf) of model and state parameters is constructed by updating the prior probability with a data likelihood function, allowing noisy dynamic observations to be integrated. However, computing the covariance matrices and the resulting probability distribution with large number of ensemble members can be computationally expensive; hence, these methods provide a practical means of sampling this probability distribution approximately with a finite number of ensemble members. If data are assimilated sequentially, the methods are referred to as ensemble filters, while in other cases, all data are assimilated simultaneously, and those methods are considered as ensemble smoothers (Chen and Oliver 2012). To further improve the convergence behavior, techniques such as ensemble randomized maximized likelihood (EnRML) have been proposed where gradients, calculated using an

adjoint system, are integrated (Oliver et al. 1996; Chen and Oliver 2012). Since the relationship between model parameters and data observations is highly non-linear, applying the updating schemes iteratively during each assimilation step has been proposed (Chen and Oliver 2012; Li and Reynolds 2009).

Ensemble Kalman filter (EnKF) is a Monte-Carlo based technique that assimilates data sequentially following the Bayesian updating framework (Evensen 2003). This technique was first proposed by Evensen (1994) and was later introduced to the petroleum engineering industry as a promising approach for handling high-dimensional history matching problems (Lorentzen et al. 2001; Aanonsen et al. 2009). The algorithm begins with generation of an ensemble of initial models (typically 40-100) that is consistent with prior knowledge of the initial state and its probability distribution (Liu and Oliver 2005). Each initial model or ensemble member is advanced to the time of the next observation using a forward model (e.g., numerical reservoir simulation), covariance and mean of model states (prior joint probability distribution) are approximated directly from the ensemble of states and used to compute the Kalman gain, which is applied to update each ensemble member. The updated ensemble provides an empirical estimate of the posterior joint probability distribution of the model states. Burgers et al. (1998) have shown that the updated ensemble obtained with EnKF approximates the theoretical posterior distribution for large ensemble size for (1) linear dynamics and measurements and (2) Gaussian prior and likelihood. Therefore, under the assumption of Gaussian statistics, EnKF updating of variables characterized by non-Gaussian statistics (e.g., multimodal distribution,

curvilinear, and channelized features) may introduce certain problems (Dovera and Della Rossa 2010). In many cases, statistics of the final updated model variables would deviate significantly from the non-Gaussian distribution exhibited by the initial model states (Zafari and Reynolds 2007).

The Gaussian statistics assumption presents an important challenge for applying EnKF to estimate facies distributions, which are commonly multimodal and spatially correlated with the underlying non-Gaussian geological features. A common strategy for working with non-Gaussian model parameters is to transform them from the original space to other space(s) in which they follow a nearly Gaussian distribution, such that EnKF can be applied in the transformed space. Different parameterization schemes such as truncated pluri-Gaussian, level set method, Gaussian mixture models (GMM), kernel principal component analysis (PCA), normal score transform, and discrete cosine transform (DCT) have been proposed in the literature with varying degrees of success (Evensen 2007; Aanonsen et al. 2009). Satisfactory results have been demonstrated with the truncated pluri-Gaussian parameterization technique for history matching of facies distribution using EnKF (Agbalaka and Oliver 2008; Liu and Oliver 2005). Agbalaka and Oliver (2008) have presented a detailed analysis of the facies proportions characterization in a three dimensional synthetic model consists of three facies. A number of researchers have implemented the level set method to update facies distribution in a channelized synthetic reservoir model with two distinct facies (Chang et al. 2010; Lorentzen et al. 2012; Moreno and Aanonsen 2011). Dovera and Della Rossa (2010) have implemented the GMM

69

to update a bimodal distribution of log-permeability values using EnKF. Kernel PCA has been used by Sarma et al. (2007, 2008) for data assimilation in a channelized reservoir model. Li et al. (2012) and Nejadi et al. (2012b) have applied normal score transform to estimate bimodal distributions of aquifer model parameters and permeability tensors in fractured reservoirs, respectively.

Discrete cosine transform (DCT), which has its roots in image processing (Jain 1988; Rao and Yip 1990), has been proposed by Jafarpour et al. (2008) to parameterize non-Gaussian distributions. They showed that spatial continuity of the channelized features can be captured effectively with a reduced set of DCT coefficients (i.e., a reduced parameter space). Although the spatial arrangement of high and low permeability regions is captured, porosity and permeability values, instead of facies indicators, were updated directly; hence, characterization of distinct facies boundaries was not addressed. Nejadi et al. (2011) introduced DCT-EWEnKF where the mean model state is calculated using a weighted scheme instead of arithmetic averaging. The updated models exhibit the correct facies boundaries, while preserving the facies proportions exhibited by the initial model states.

This chapter attempts to address two particular issues pertinent to applying EnKF for history matching of facies distributions. The first issue is the capability of the algorithm to represent the non-Gaussian statistics exhibited by the initial model states. Prior to the data assimilation stage, an ensemble of initial models that is consistent with prior knowledge of the initial state and its probability distribution inferred from static geological data is generated. Although

70

description of geological features such as facies distributions (proportions and spatial patterns) and channel properties (e.g., orientation, widths, and amplitudes) can be derived from various sources of static geologic information, there is no mechanism within EnKF that would allow explicit conditioning of the posterior distribution, during the calculation of Kalman gain or the updating step, to the reference statistics (e.g., histogram and semi-variogram or covariance) inferred from this static information.

The second issue is the ability to maintain diversity within ensemble during updating. Limited data observability and finite ensemble size often result in a corrupted estimate of the cross covariance. Due to limitation in data observability and insufficient degrees of freedom, production history, particularly at the early stages, contains inadequate information to properly characterize the probability distribution of all model parameters. As a result, the variability across the ensemble would diminish substantially after a few EnKF assimilation steps (Jafarpour et al. 2011); realizations with poor initial estimates would converge towards a single ensemble member after major updating to match the early production history. In addition, an ensemble of finite size is typically implemented in most practical applications of EnKF; the Kalman gain calculated from an ensemble of reduced variability could cause the ensemble to collapse after successive updates (Furrer and Bengtsson 2007; Myrseth et al. 2012). In order to avoid the spurious and unrealistic correlations between the state vector and predictions, localization techniques have been introduced to constrain the updates of observations (Chen and Oliver 2009, 2011; Emerick and

Reynolds 2010). Ensemble adjustment, by means of covariance inflation, has also been introduced by Anderson (2001) to maintain the ensemble variability. These authors have implemented various distance-based functions for covariance localization. Arroyo et al. (2008) described a flow-based localization method of conditioning the covariance matrix using information derived from streamline trajectories. Only grid blocks that are sensitive to production response are used in the covariance calculation, eliminating the influences of spurious correlations. The authors demonstrated that spatial continuity can be maintained. In all these approaches, the ensemble is advanced along the assimilation process without re-sampling. Other authors such as Emerick and Reynolds (2012), however, have proposed repeating the history matching procedure by integrating the Markov Chain Monte Carlo algorithm along with EnKF to generate an accurate sampling of the posterior distribution. In their work, EnKF is applied multiple times, and a number of Markov chains are generated from the final ensembles.

The motivation of this chapter is to present a modified approach to Nejadi et al. (2012a, 2014) and propose improvement on the two aforementioned deficiencies when applying EnKF for facies distribution modeling: to maintain the reference statistics of the model parameters and to improve ensemble variability. This is achieved within the conventional framework of EnKF using a re-sampling step. The assimilated ensemble members are used to sample an improved ensemble that also enhances the diversity among ensemble members at an intermediate stage during the assimilation process when data mismatch have reduced substantially but the ensemble variance is still high. The new improved ensemble

incorporates the dynamic model-updating information of the initial steps and is consistent with the static geological data; it is then subjected to the forecast model (i.e., reservoir flow modeling) from the beginning to the last update step prior to re-sampling, and updating would resume for the remaining production history using EnKF. If necessary, covariance regularization techniques such as localization and inflation can also be implemented at each subsequent EnKF update step.

In the examples presented in this chapter, the re-sampling is performed only once. The selection criteria based on data mismatch and ensemble variance for when re-sampling should be performed is discussed. Although performing flow simulations over the re-sampled ensemble from the beginning reduces computational efficiency, the additional cost could be easily justified by the important potential for applying the Kalman filter to non-Gaussian model parameters with non-linear process dynamics (Gu and Oliver 2007; Li and Reynolds 2009). In the proposed procedure, the re-sampling step allows the reference histogram and spatial correlation of the model parameters derived from static geologic information to be re-established in the updating process. Conventional implementation of EnKF continues to update this re-sampled ensemble based on the dynamic information available, without altering the posterior probability distribution of the model states through any other explicit conditioning.

This chapter is organized as follows: first, the formulation of the re-sampling procedure to generate new ensemble members is detailed. Next, results obtained

from the application of the proposed methodologies in two synthetic reservoir models are presented. The first model has three facies and the second model is a channelized reservoir with two facies. At the end, conclusions and remarks drawn from the case studies are discussed.

3.2 Methodology

In this section, various components of the implementation are discussed in detail. Specific modifications to the conventional EnKF procedure are highlighted.

3.2.1 Forecast step

The forecast model (g) is separately applied to each ensemble member, advancing the model and state variables to the next time step. Model equations pertinent to the multiphase fluid flow in petroleum reservoirs are described by material balance, momentum balance, phase behavior descriptions, and numerous auxiliary relations (Oliver et al. 2008). This non-linear system of differential equations is solved numerically using methods such as finite difference or finite element

$$\begin{bmatrix} u_k^p \\ d_k^p \end{bmatrix} = g \begin{bmatrix} m_{k-1}^a \\ u_{k-1}^a \end{bmatrix}.$$

3-1

In the above equation, m and u denote the model and state variables, respectively, and d_k represents the predicted production data at k^{th} step. The superscripts, p and a, specify the predicted and analyzed states, respectively. In this study the commercial reservoir simulation software Eclipse black-oil (2011) was used as the forecast model.

3.2.2 Analysis step

Prior to any updating, an initial ensemble of realizations is generated using Monte Carlo simulation. These realizations or ensemble members are created based on the available geological data, capturing the uncertainty inferred from static geologic information or a priori knowledge about the model variables. The ensemble of state vectors is denoted in vector form as

$$\psi_k = \left\{ y_k^1, y_k^2, \dots y_k^{N_e} \right\},$$

3-2

where N_e is the total number of ensemble members and $y_k^j, j = 1, \dots, N_e$ is state vector corresponding to the jth ensemble member at k^{th} assimilation step. State vector should consist of model parameters (m), such as porosity, permeability, or transformed variables of facies indicators at all grid locations; time-dependent state variables (u), such as pressure and saturations at all grid locations; and observed production data (d) including production and pressure measurements recorded at well locations as a function of time

$$y_k = \left\{ \begin{matrix} m_k \\ u_k \\ d_k \end{matrix} \right\}.$$

3-3

EnKF propagates the ensemble of initial state vectors along time to assimilate data in a two-step approach. As discussed in the previous section, in the first (forecast) step, each ensemble member is subjected to the forward model g

$$\begin{bmatrix} u_k^{p,j} \\ d_k^{p,j} \end{bmatrix} = g \begin{bmatrix} m_{k-1}^{a,j} \\ u_{k-1}^{a,j} \end{bmatrix}, \quad j = 1, \dots, N_e$$

$$\text{3-4}$$

In the second (analysis) step, the state vectors are updated with the observation data d_k^j using following equation

$$y_k^{a,j} = y_k^{p,j} + K_k \left(d_k^j - H y_k^{p,j} \right).$$

$$\text{3-5}$$

The forecast and analysis steps are repeated sequentially in time whenever new observations are available. In Equation **3-5**, H is the observation operator, which represents the relationship between the state vector and the observation vector

$$H = [0 \mid I],$$

$$\text{3-6}$$

where d_k^j is the summation of the observed production data at the kth assimilation step $(d_{obs,k})$ and ε_k^j, a vector of measurement errors such that

$$E\left[\varepsilon \varepsilon^T \right] = C_D$$

$$d_k^j = d_{obs,k} + \varepsilon_k^j.$$

$$\text{3-7}$$

K_k is the Kalman gain at the kth assimilation step

$$K_k = C_{y_k^p} H^T \left(H C_{y_k^p} H^T + C_{Dk} \right)^{-1},$$

$$\text{3-8}$$

where $C_{y_k^p}$ is the state cross covariance matrix and C_{Dk} is the measurement error covariance matrix. The cross covariance matrix is estimated from the ensemble of state vectors as

$$C_{y_k^p} \approx \frac{1}{N_e - 1} \left\{ \left(y_k^p - \overline{y_k^p} \right) \left(y_k^p - \overline{y_k^p} \right)^T \right\}$$,   3-9

where $\overline{y_k^p}$ is the mean state vector.

3.2.3 Discrete Cosine parameterization of non-Gaussian facies indicators for continuous reservoir model updating in ensemble Kalman filter

In history matching with Ensemble Kalman filter and discrete cosine parameterization (DCT-EnKF), DCT is performed to parameterize the non-Gaussian facies indicators into coefficients of the retained cosine basis functions. These coefficients are incorporated into the state vector as model parameters (m_k) in Equation 3-3, and they are updated subsequently in the EnKF procedure. At the end of the analysis step, they are transformed back into facies indicator values via inverse DCT (Jafarpour et al. 2008). In order to reduce the size of model parameter space and retain the ability to capture the large-scale spatial continuity of different facies, a reduced set of DCT coefficients, which comprises 20% of all coefficients with the highest absolute values, are included in the state vectors. Jafarpour et al. (2008) discussed a number of approaches for selecting a reduced set of coefficients. This work adopted one of their approaches and performed a sensitivity analysis to determine the optimal size of the reduced parameter set. It is observed that when too few coefficients are used, correct spatial continuity cannot be captured satisfactorily; while if all coefficients are used, the resultant facies distribution exhibits unphysical short-

scale variability, requiring implementation of smoothing algorithms for post-processing (Nejadi et al. 2012a). As a result, it is concluded that 20% of all coefficients would be the optimal choice.

3.2.4 Generation of initial ensemble

Initial ensemble of facies distribution is generated by Sequential Indicator Simulation (SIS) (Deutsch and Journel 1997), conditioned to facies observations at the well locations. Spatial statistics including histogram and semi-variogram model $\gamma(l)$ are inferred from well or other geologic information. Considering a random function $Z(x), x \in R^{N_b}$ in a discretized system with N_b grid cells, the cumulative probability distribution of $Z(x)$ together with the indicator transform of $I(Z(x);z)$ are given by Equation 3-10

$$I(Z(x);z) = \begin{cases} 0 & if \quad Z(x) \le z \\ 1 & if \quad Z(x) > z \end{cases}$$

$$F_{Z(x)}(z) = P\{Z(x) \le z\} = E\{I(Z(x);z)\}, \forall x \in R^{N_b} \quad . \qquad \text{3-10}$$

A set of conditioning hard data Λ_0: $h_{N_w} = \{h_w, w = 1,...,N_w\}$ represents the measurement available at N_w well locations. The conditional simulation in the grid system with N_b grid cells is based on sampling from the N-variate distribution that is conditioned upon Λ_0

$$F_{Z(x_1),...,Z(x_{N_b})}\left(z_1,...,z_{N_b} | \Lambda_0\right) = P\left\{Z(x_1) \le z_1,...,Z(x_{N_b}) \le z_{N_b} | \Lambda_0\right\} . \qquad \text{3-11}$$

3.2.5 Ensemble re-sampling with probability weighting

In this section, a novel re-sampling procedure is proposed to address two common issues associated with conventional EnKF implementation: (1) incorporating the non-Gaussian statistics inferred from static geologic information, as exhibited by the initial model states and (2) maintaining diversity among ensemble members during updating. In this approach, at an intermediate stage during the assimilation process when data mismatch have reduced substantially but the ensemble variance is still high, an improved ensemble is generated following a re-sampling procedure. A set of re-sampling points are selected from among the model parameters (facies indicators at all grid blocks) where variance reduction is most significant after the initial updating. Details of this selection criterion will be discussed in next section. Individual values of the model parameters at these re-sampling points are obtained by sampling from the corresponding probability distributions calculated using the updated ensemble just prior to re-sampling. Along with the hard data, they are regarded as conditioning data for generating a new ensemble via conditional sequential simulation. The new ensemble incorporates the model-updating information of the initial steps and is consistent with the static geological data; it is then subjected to the forward model from the beginning to the last update step prior to re-sampling, and updating would resume for the remaining production history using EnKF.

New ensemble members are again generated by performing SIS, but in contrast to the initial ensemble, they are conditioned to both hard data h_{Nw} available at the

79

well locations and facies indicator values h_{rp} sampled from probability distributions derived from the updated ensemble members at selected re-sampling points. The set of conditioning data for the simulation is represented by $\Lambda_{rp} : \{h_{Nw}, h_{rp}\}$, and the corresponding multivariate probability distribution is defined as

$$F_{Z(x_1),\dots,Z(x_{N_b})}\left(z_1,\dots,z_{N_b}\big|\Lambda_{rp}\right) = P\left\{Z(x_1) \le z_1,\dots,Z(x_{N_b}) \le z_{N_b}\big|\Lambda_{rp}\right\}. \qquad 3\text{-}12$$

The resulting realizations will be consistent with both the initial dynamic data obtained from the previous EnKF assimilation steps. Statistics from static geologic information, which is difficult to retain during traditional EnKF update, would be re-established in the new or re-sampled ensemble.

A practical consideration is the specification of the location and number of re-sampling points. RamaRoa et al. (1995) proposed placing the conditioning points (referred as pilot points in their studies) in high sensitivity zones where they have the highest potential to elicit a reduction in the objective function and correctly infer the underlying heterogeneities. In the next section, details regarding the selection and number of re-sampling point selection are presented. At the end of this methodology section, the criterion for when re-sampling should be carried out will also be discussed.

Selection of re-sampling points and generation of new ensemble members

This work proposes a procedure by which the re-sampling points are selected dynamically (instead of being determined a priori) among the model parameters that have experienced the largest reduction in ensemble variance. Given that

EnKF is a variance minimizing technique, model parameters that exhibit a multivariate link with the observations would undergo a variance reduction (van Leeuwen 2003). This observation implies that maximum reduction in the ensemble spread of model parameters is realized where the coefficients of the sensitivity matrix $\partial g(m)/\partial m$ are the greatest. In other words, these model parameters have high impact on the production performance of the reservoir and should be selected as potential re-sampling points. As discussed in the previous section, approximation with an ensemble of finite size and limited data observability necessitate covariance localization to mitigate issues arising from spurious correlation and to maintain ensemble diversity. In a simple univariate sense, the ensemble variance is defined as

$$\sigma_z^2 = \sum_{j=1}^{Ne} \left(z_k^{j,p} - \bar{z}_k^p \right)^2 \Big/ \left(N_e - 1 \right)$$
, 3-13

where $\bar{z}_k^p$ is the ensemble average of $z_k^{j,p}$. For multiple facies models (multinomial distribution) variance is defined as $Var(z_i)=np_i(1-p_i)$, where n is the number of ensemble members (N_e) and p_i is the probability of a specific facies at grid block location k.

The idea is to limit the parameter updating to specific regions of the model that are most sensitivity to the observation data. As discussed in the literature review, techniques of flow-based localization and re-sampling are useful to alleviate issues related to spurious correlation. Re-sampling is particularly useful for re-establishing in the ensemble the reference histogram and variogram statistics inferred from the original static data.

Denoting $f_{Z(x)}(z_i)$ as the probability of occurrence of a particular facies z_i at a location x and the total number of distinct facies as n_f, at the end of each assimilation step, the cumulative density function $F_{Z(x)}(z_i)$ at x can be constructed by

$$F_{Z(x)}(z_i) = \sum_{t \leq i} f_{Z(x)}(z_t) \qquad z_i = z_1, z_2, ..., z_{n_f}$$

3-14

Re-sampling points are selected where the uncertainty in facies distribution has been reduced significantly and the probability of the occurrence of the most probable facies is high. This criterion is equivalent to assessing whether the cumulative probability distribution calculated with Equation 3-14 would resemble that of a Heaviside step function. It could be implemented by assigning to each location a sampling weight, whose value reflects the similarity of the probability function to that of a Heaviside step function. The sampling weight is assigned following the concept of importance weights in particle filtering techniques, where ensemble members (particles) are weighted on the basis of likelihood function. Emerick and Reynolds (2012) suggested a weighting function that is proportional to exp($-O$(m)), where O(m) is a user-defined objective function. In this work, the objective function at each location is formulated as $\sum_{i=1}^{n_f} \left| H(z^*) - F_{Z(x)}(z_i) \right|$, where $H(z^*)$ represents the Heaviside step function: its value is zero if $z_i < z^*$ (where z^* is the most probable facies) and its value equals one if $z_i \geq z^*$. Locations with high weights will be selected preferentially as re-sampling points.

Specific values of model parameters (i.e., facies indicators) at the re-sampling points are sampled from the probability distributions $F_{Z(x_{rp})}(z_i), i = 1...n_f$, where x_{rp} is the location x of a certain re-sampling point denoted by the subscript rp. In order to ensure consistency with conditioning data observed at the well locations, the hard data (h_{Nw}) is also included in the set of re-sampling points. These data are used in a sequential simulation algorithm to generate a new ensemble of model parameters.

To ensure the aforementioned sampling procedure is capturing the uncertainty represented by $F_{Z(x_{rp})}(z_i), i = 1...n_f$, several sets of re-sampling points are selected as conditioning data and different sets of model parameters are generated. The new sets of model parameters are combined, and they constitute the re-sampled ensemble. This avoids introducing artificial bias in the conditioning data and under-representing the variability among members of the ensuing re-sampled ensemble. Ensemble spread, which quantifies the uncertainty of the parameter space (Chen and Zhang 2006), can be used to determine the optimum number of sets of re-sampling points

$$ES(z) = \left[\frac{1}{N_b} \sum_{i=1}^{N_b} \sigma_{zi}^2 \right]^{1/2},$$

3-15

where σ_{zi}^2 is the value of σ_z^2 in Equation 3-13, at grid block i. As the number of sets increases, the ensemble spread gradually increases to a maximum value and stabilizes thereafter. Figure 3-1(a) shows the ensemble spread as a function of the number of re-sampling sets for Example A in the next section. Following the

previous notation, value of the ensemble spread represents the global expectation of the variance of $F_{Z(x)}(z)$ evaluated over all locations of x. The ensemble spread should also be interpreted as a measure of variability exhibited by members of the re-sampled ensemble. Therefore, by utilizing several sets of re-sampling points and constructing multiple realizations corresponding to each set, the uncertainty associated with the model parameter space can be assessed. According to Figure 3-1(a), improvement in ensemble variance with more than 5 sets of re-sampling points is inconspicuous; hence, the optimal number of sets of re-sampling points should be 5.

The forecast model is applied to the re-sampled ensemble from beginning to the last update step prior to re-sampling to generate consistent state variables u^r and the predictions of production data d^r. The new ensemble of state vectors is constructed

$$\Psi^r = \left\{ y^{r1}, y^{r2}, \ldots, y^{rN_e^r} \right\},$$ 3-16

where Ψ^r is the ensemble of new (re-sampled) state vectors. The superscript r denotes the new vector after re-sampling and N_e^r is the number of members in the re-sampled ensemble.

Number of re-sampling points

The number of re-sampling points should be selected depending on the nature of the problem. Factors such as size of reservoir model, well pattern, production mechanism, reservoir heterogeneity and other parameters should be taken into consideration to determine the number of re-sampling points. The number of re-

sampling points may be optimized such that they are large enough to capture the adjustments made to model parameters during the early production data assimilation. On the other hand, if all model parameters are selected as re-sampling points and used as conditioning data for constructing the re-sampled ensemble, uncertainty in the model parameters would be under-represented, causing an early ensemble collapse. This also implies that the statistics (i.e., experimental semi-variogram or average channel width) computed for the re-sampled ensemble members would not maintain those of the reference statistics. As a result, the number of re-sampling points should also be small enough to maintain variability among re-sampled ensemble members and allow spatial uncertainty to be captured adequately.

A sensitivity analysis was performed to estimate the optimum number of re-sampling points, where ensembles with different number of re-sampling points were generated, and the experimental semi-variogram of all ensemble members were calculated. It is anticipated that with more re-sampling points, deviation of the experimental semi-variogram from the reference semi-variogram would increase. Therefore, one should select as many re-sampling points as possible, while minimizing the mismatch between the experimental semi-variogram of the re-sampled ensemble members and the reference variogram. This mismatch can be quantified as a weighted nonlinear least squares error (Cressie 1985)

$$R = \sum_{i=1}^{n} \frac{N_{li}}{\gamma(l_i)^2} \left[\gamma^*(l_i) - \gamma(l_i) \right]^2$$

,   3-17

where $2\gamma(l_i)$ is the experimental variogram at lag distance l_i evaluated for a particular re-sampled ensemble member, and γ^* represents the reference variogram. Equation 3-17 is used to calculate the mismatch exhibited by an individual member, and Figure 3-1(b) depicts its average value over all re-sampled ensemble members as a function of the number of re-sampling points. This average error begins to rise as the number of re-sampling points increases beyond 20 percent. Therefore, the optimum number of re-sampling points has been estimated to be 20 percent of the total number of grid blocks in the cases studies.

To validate the reproduction of curvilinear geological patterns between multiple facies based on multiple point statistics, the concept of distribution of runs (Mood 1940) is implemented. Consecutive sequences of a specific category are counted. The difference in distribution of runs between the re-sampled realizations and the training image is calculated according to Boisvert et al. (2008)

$$R = \sum_{l=1}^{n} \left| f_l^* - f_l \right|$$

,   3-18

where f_l is the cumulative frequency of facies indicators for a run of length l in a specific direction for a particular re-sampled ensemble member, f_l^* is the corresponding cumulative frequency from the training image, n is the maximum length of runs. Equation 3-18 is used to calculate the mismatch exhibited by an

individual member. As in the variogram case, the average error computed over all re-sampled ensemble members is plotted against the number of re-sampling points; approximately 20 percent of the total number of grid blocks is determined to be the optimal number of re-sampling points in the cases studies.

Re-sampling step

The proper choice of number of assimilation step(s) or time at which to perform re-sampling is discussed here. As the number of assimilation steps increases, both the objective function and ensemble variance would diminish. Since re-sampling points are chosen at locations that are most sensitive to the observation data, it implies that re-sampling should take place after the objective function has experienced a moderate decline. However, to avoid ensemble collapse and maintain diversity among the re-sampled ensemble members, this re-sampling should be performed while the ensemble variance is still high. The objective function is defined as the dynamic data mismatch

$$O(m) = [g(m) - d_{obs}]^T C_D^{-1} [g(m) - d_{obs}]_.$$

3-19

At the beginning of the assimilation procedure, the mismatch is high, reflecting the uncertainty in the initial ensemble and the respective proximity of each member to the true dynamic responses. As the assimilation procedure continues, the objective function $O(m)$ decreases. A major portion of the ensemble uncertainty is corrected after the first few assimilation steps (Jafarpour et al. 2011).

As the updating continues, ensemble members with poorly estimated initial model states would not experience further reduction in the objective function, while the overall ensemble diversity incessantly decreases after each and every update step. Re-sampling carried out at that point would yield an underestimation of the variance of the posterior probability distribution of the model parameters (Emerick and Reynolds 2012). The proper re-sampling step can be quantitatively determined by plotting the normalized ensemble variance of model states and the data mismatch ($O(m)$) simultaneously as a function of assimilation steps. In the case studies, the re-sampling was carried out when the ensemble variance had decreased substantially and stabilized thereafter; this point typically coincides with a significant reduction in data mismatch after the first few updating steps. Further details are provided in the next section, where two examples are presented. Figure 3-2 summarizes all the prescribed steps.

3.3 Implementation of Ensemble Kalman Filter with Re-sampling

The proposed re-sampling point simulation technique has been implemented for history matching of two synthetic reservoir models: one with three distinct facies and another channelized model with two facies. Facies distribution in space is the unknown model parameter in both examples. The objective is to update the facies indicators by means of integrating observed data, while honoring conditioning data at well locations and preserving the reference statistics (facies proportions and spatial relation) derived from static geologic information. As

described in the previous section, the proposed approach incorporates into the conventional EnKF framework a new re-sampling step that consists of (1) selection of re-sampling points, (2) generation of new ensemble members, and (3) updating of state vectors by applying the forecast model from the beginning.

3.3.1 Model Description – Example A

A synthetic two-dimensional reservoir model was considered. The model has three distinct facies namely sand (S), fine sand (FS) and shale. The reference facies proportions were 45% S, 37% FS, and 18% shale. Permeability within each facies is assumed to be uniform with constant porosity throughout the model.

$$\phi = 30\% \qquad \text{all facies}$$

$$k(\text{md}) = \begin{cases} 2000 & \text{for Sand} \\ 500 & \text{for Fine Sand} \\ 50 & \text{for Shale} \end{cases}$$

Facies distribution for the reference (benchmark) model was generated via unconditional SIS, as implemented in the GSLib (Deutsch and Journel 1997). An exponential variogram model with the maximum and minimum correlation lengths of 1370 m and 518 m, respectively, along an azimuth angle of 135$^{\text{o}}$ (counter clockwise from east) is used. Model contained 21×21 grid blocks along the x- and y- directions. The dimensions of the model are 1600 m × 1600 m × 6 m. The production mechanism was waterflooding; four injectors and one producer operating at constant bottom-hole pressures, were placed in a five spot pattern, along with four additional observation wells, as shown in Figure 3-3(a).

Figure 3-4(a) shows the facies distribution of the reference model and Table 3-1 summarizes the parameters of the dynamic model.

Hard data was collected at all 9 well locations, which include facies classifications. Observation measurements (collected on a monthly basis) consist of oil rate and water cut at the producer, water injection rate of the injectors, and bottom-hole pressure of the observation wells, were obtained by subjecting the reference model to the forward model (reservoir simulator).

3.3.2 Base case – Example A

A case was set up to establish a basis for comparison by examining the conventional DCT-EnKF method and its performance for updating multiple facies distributions. An initial ensemble of 100 facies distribution was generated via conditional SIS. Facies observations at the well locations (producer, injectors, and observation wells – 9 hard data) were used as the conditioning data. The variogram parameters were same as the true case. Facies proportions of the initial ensemble members were within the range of 5% of the reference distribution. The production schedule is defined for 32 years. The first 10 years of production data were used for history matching (20 assimilation steps of every 6 months). Forward modeling with the history-matched ensemble was carried out for the remaining 22 years to evaluate the prediction capability. In Figure 3-5, the production and injection data before and after history match are shown, while Table 3-2 summarizes the root mean square error (RMSE) of the production match for the entire period of 32 years based on the Nash-Sutcliffe model efficiency (Nash and Sutcliffe 1970)

90

$$RMSE^j = 1 - \frac{\sum_{k=1}^{N_k}(d_{obs,k} - d_k^{p,j})^2}{\sum_{k=1}^{N_k}(d_{obs,k} - \overline{d_{obs}})^2},$$

3-20

$$RMSE = \left[\sum_{j=1}^{N_e} RMSE^j\right] \Big/ N_e,$$

3-21

where $d_{obs,k}$ is observed values; the modeled values for ensemble member j at time step k is denoted by $d_k^{p,j}$, and N_k is the total number measurement time steps. The RMSE value derived from this formula can range from $-\infty$ to 1, where a value of 1 (RMSE = 1) corresponds to a perfect match. RSME values of oil production rate, produced water cut, water injection rate, and bottom hole pressures at observation wells are computed using Equation 3-20 and Equation 3-21.

Figure 3-6 compares the average permeability map of the initial and assimilated ensemble. Figure 3-7 shows the permeability map of a randomly selected ensemble member before and after update. It is apparent that despite production history has been matched; spatial continuity of facies distribution in the updated ensemble is different from the reference case.

3.3.3 Re-Sampling – Example A

In this example, the proposed re-sampling scheme was applied after eighteen months of production history matching (i.e., 3rd assimilation step of Base Case). Figure 3-8 depicts the calculated objective function, Equation 3-19, after different assimilation steps for the Base Case (Example A). After a few

assimilation steps, the value has decreased from its initial value but remains relatively constant until water breakthrough. As discussed previously, major portion of the ensemble uncertainty has been corrected after the few initial assimilation steps, which is in agreement with observations made by Jafarpour et al. (2011). Also shown in Figure 3-8 is the normalized ensemble variance of model parameters at a few selected assimilation steps. Between the 3rd and 13th step, there is no observable decrease in the value of the objective function, indicating that EnKF assimilation has not improved the quality of the history match; however, the ensemble variance continues to diminish. Given that the proper re-sampling step is where a substantial decrease in the objective function is observed, without compromising the ensemble variability, it is concluded that for this example, re-sampling should be carried out after the 3rd assimilation step. Using the updated ensemble members after three assimilation steps, five different sets of re-sampling points were selected. Each set of re-sampling points, along with the observed facies classifications at the well locations, were used to generate a new re-sampled ensemble with 20 members. A total of 100 re-sampled ensemble members were further assimilated with DCT-EnKF. Figure 3-9 shows the average permeability map of the five re-sampled ensembles along with the corresponding re-sampling point locations. This new ensemble of 100 members is then subjected to the forward model from the beginning to the last update step prior to re-sampling (i.e., 3rd assimilation step). EnKF updating is resumed to assimilate the rest of production history.

Figure 3-10 compares the production and injection profiles of the re-sampled ensemble before and after resuming EnKF updating. The corresponding RMSE values of history match results as summarized in Table 3-2 are much closer to unity than the Base Case. Average maps of permeability distribution before and after update are shown in Figure 3-11. As compared to the Base Case (Figure 3-6), the resulting permeability map has shown significant improvement in terms of exhibiting similar spatial characteristics as depicted by the reference distribution. As an example, Figure 3-12 also shows the permeability map of a randomly selected ensemble member before and after update.

For the same randomly selected ensemble member, Figure 3-13 compares the normalized experimental semi-variogram calculated along the minor and major directions of anisotropy for the reference model, base case and the re-sampling case. It is clear that only the re-sampling scheme could ensure reproduction of the reference variogram. Similar improvement is observed for other ensemble members. Figure 3-14 compares the average facies proportions of the updated ensemble for different cases. For the base case, where re-sampling was not performed, the reference facies proportions could not be reproduced. On the other hand, the facies proportions with the re-sampling scheme resemble much closely to the reference statistics. The obtained results clearly demonstrate the improved performance of the proposed re-sampling technique.

Value of the ensemble spread calculated using Equation 3-15, decreases from 1.57 for the initial ensemble to 1.05 at the 3rd assimilation step prior to re-sampling. This value rises to 1.20 immediately after re-sampling and decreases

again to a final value of 0.45 at the last assimilation step (end of history matching period). It is interesting to note that this final value would be 0.39 for the Base Case. This observation indicates that re-sampling has improved the ensemble variability, as evidenced by the increase in ensemble spread.

3.3.4 Model Description – Example B

In this example, a synthetic model (100×80 grid blocks along the x- and y-directions) with a low permeability background (Fine Sand) and high permeability channels (Sand) was considered. The dimensions of the model are 1000 m × 800 m × 10 m. The reference facies proportions were 69% FS and 31% S. Permeability within each facies is assumed to be uniform with constant porosity throughout the model.

$$\phi = 20\% \qquad \text{all facies}$$
$$k(\text{md}) = \begin{cases} 1000 & \text{for Sand} \\ 50 & \text{for Fine Sand} \end{cases}$$

Facies distribution for the reference model was reconstructed from the channelized model permeability distribution as presented in Li et al. 2012.

A horizontal water injector and a horizontal oil producer, both equipped with inflow control valves (ICV) and operated at constant bottom-hole pressure, were placed in the model. Figure 3-3(b) depicts the well configurations. Figure 3-4(b) shows the facies distribution of the reference model and Table 3-1 summarizes the parameters of the dynamic models.

Hard data was collected at well locations, which include facies classification, porosity, and permeability values. Observation measurements (collected on a

94

monthly basis over a period of 18 months) consist of water injection rate of the injector, oil rate and water cut at the producer were obtained by subjecting the reference model to the forward model (reservoir simulator).

3.3.5 Base case – Example B

The initial ensemble was generated by Single Normal Equation Simulation (SNESim), as implemented in SGeMS (Remy et al. 2009), using the training image shown in Figure 3-15. Figure 3-16 compares the production and injection profiles before and after history match, and the corresponding RMSE values of history match results are summarized in Table 3-2. RMSE values of oil production rate, production water cut, and water injection rate data of the injectors are computed using Equation 3-20 and Equation 3-21.

Figure 3-17 compares the average permeability maps of the initial ensemble and the updated ensemble after 9 assimilation steps. Nine months of production data were used for history matching (9 assimilation steps, every month). Figure 3-18 shows the permeability map of a randomly selected ensemble member before and after update. It is apparent that despite production history has been matched, spatial continuity of facies distribution in the updated realizations are different from the reference case.

3.3.6 Re-Sampling – Example B

Similar to Figure 3-8, the calculated objective function and normalized ensemble variance of model parameters are plotted in Figure 3-19. Between the 5th and 9th step, the value of the objective function remains constant, implying that EnKF

assimilation has not improved the quality of the history match; however, the ensemble variance continues to diminish. Therefore, it is concluded that for Example B, re-sampling should be carried out after the 5th assimilation step.

Ensemble spread was calculated using Equation 3-15, and the optimum number of re-sampled ensemble sets was estimated to be five. Each set of re-sampling points, along with the observed petrophysical properties along the horizontal wells, were used to generate 20 new ensemble members. The combined ensemble (a total of 100 members) was further assimilated with EnKF.

Figure 3-20 compares the production and injection profiles of the re-sampled ensemble before and after resuming EnKF updating. The corresponding RMSE values of history match results are summarized in Table 3-2. Average permeability distributions of the initial and final updated ensemble are shown in Figure 3-21. Figure 3-22 shows the permeability map of a randomly selected ensemble member before and after update. The updated ensemble is acceptable in terms of RMSE values and the resulting permeability map has shown significant improvement, as compared to the Base Case (Figure 3-17), in terms of exhibiting similar spatial characteristics as depicted in the training image and the reference case. Figure 3-23 compares the average facies proportions of the updated ensemble for different cases of Example B; the re-sampling scheme is, again, performing better than the Base Case in reproduction of reference histogram.

3.4 Results and Discussion

Different parameters were considered in the examples to assess the performance of the methodologies. Firstly, the RMSE value was used to measure the quality of the observed data match. Secondly, experimental semi-variogram of the reference case and the updated ensembles were compared. Finally, the facies proportions of the history matched models and the reference model were compared to verify if the correct proportions were maintained in the posterior ensemble.

In each of the examples, it is illustrated through the comparison with the base cases that although production history match has been achieved after several assimilation steps, spatial continuity of geologic features are no longer maintained with conventional EnKF. This is evident by comparing Figure 3-6(b) and Figure 3-7(b) with Figure 3-11(b) and Figure 3-12(b). The same observation can be made in the channelized model, where facies continuity is lost after several update steps, as evidenced in Figure 3-17(b) and Figure 3-18(b). Figure 3-21(b) illustrates how the re-sampling scheme has successfully re-established this spatial continuity in the final updated ensemble.

Prior to any updating, few hard data are typically available to generate the initial ensemble. Many researchers have reported the issue of ensemble collapse due to a poor initial ensemble. The ensemble typically experiences major updates during the first few steps as well as right after an additional observation data type (e.g., water cut) becomes available. This comment is illustrated in Figure 3-8 and Figure 3-19, which show the objective function (Equation 3-19)

value decreases with successive update steps up to the 13th update step in example A and 9th update step in example B. Difference in breakthrough behavior among different ensemble members causes the objective function to increase sharply. Thereafter the model parameters are updated in accordance to the water cut measurements.

3.5 Conclusion

The new re-sampling-point formulation re-samples a new ensemble of model parameters, which has incorporated important model updating information derived from early production data. Model parameters having the highest impact on the production performance are selected as re-sampling points, which capture updating information derived from previous assimilation steps to be incorporated in the new ensemble. Besides, the re-sampled members reproduce the reference statistics regarding the proportions and spatial continuity of different facies. The re-sampling aids the filter to maintain the diversity among ensemble members and avoid underestimation of the uncertainty in the posterior probability distribution.

In order to ensure consistency among state and model variables, after every re-sampling, the entire ensemble is subject to the forward model from the beginning until the last EnKF update step. This will certainly incur additional computational efforts. However, the additional costs can be justified by the improvement in terms of reference statistics reproduction and uncertainty estimation in the posterior probability distribution.

The proposed re-sampling procedure results in lower root mean square error of production data match. At the same time, the facies proportion and variogram of the assimilated models are more consistent with the reference distribution. The technique presents a promising potential in characterization of non-Gaussian model parameters using diverse data sources.

Table 3-1 Parameters of the dynamic model for case studies.

Example-A	
Model Dimensions	21×21×1 grids in X, Y and Z directions
Grid Dimensions	76 × 76 × 6 m (250 × 250 × 20 ft)
Reservoir Depth	914 m (3000 ft)
Porosity	30% constant
Permeability Range	50, 500 and 2000 md
Operating Bottom-hole Pressure	Producer: 2300 psi Injectors: 2500 psi
Noise Standard Deviation	Producer: 50 STBPD Injectors: 20 STBPD
Initial Reservoir Pressure	2500 psi
Total Generated History	32 years (10 years was used for history matching)
Example-B	
Model Dimensions	100×80×1 grids in X, Y and Z directions
Grid Dimensions	10 × 10 × 10 m (33 × 33 × 33 ft)
Reservoir Depth	914 m (3000 ft)
Porosity	20% constant
Permeability Range	50 and 1000 md
Operating Conditions	
Producers	BHP= 2800 psi
Injectors	BHP= 3200 psi
Noise Standard Deviation	Producers: 20 STBPD Injectors: 20 STBPD
Initial Reservoir Pressure	3000 psi
Total Generated History	18 months (9 months was used for history matching)

Relative Permeability and Capillary Pressure Curves			
Sw	Krw	Kro	Pcow
0.2	0	1	7
0.3	0.07	0.4	4
0.4	0.15	0.125	3
0.5	0.24	0.649	2.5
0.6	0.33	0.0048	2
0.8	1	0	0

Table 3-2 Average RMSE values for the entire production period calculated using Equation 3-20 and Equation 3-21, for the initial and history-matched ensemble. For both examples, the initial ensemble for the Base Case was generated using only conditioning data at the well locations, whereas for Case-RP, the initial ensemble refers to the re-sampled ensemble generated using the proposed re-sampling technique. A value of 1 (RMSE = 1) corresponds to a perfect match.

Example-A	Producer		Injector				Observation			
	Oil rate	Water cut	Water injection rate				Pressure			
Well NO.	1	1	1	2	3	4	1	2	3	4
Base Case - Initial Ensemble	0.4974	0.3821	0.3652	0.4112	0.552	0.3971	0.2986	0.2853	0.3689	0.2988
Base Case - After History Matching	0.8267	0.8563	0.7410	0.8138	0.7724	0.7958	0.7222	0.799	0.7798	0.7460
Case-RP - Initial Re-sampled Ensemble	0.6025	0.6780	0.6325	0.7122	0.6560	0.7076	0.6023	0.6566	0.6750	0.6433
Case-RP - After History-Matching	0.9145	0.8672	0.9381	0.9762	0.9426	0.9565	0.8948	0.9622	0.9463	0.8797

Example-B	Producers		Injectors
Wells	Oil rate	Water cut	Water injection rate
Base Case - Initial Ensemble	0.6432	0.3025	0.4129
Base Case - After History Matching	0.9167	0.5041	0.7489
Case-RP - Initial Re-sampled Ensemble	0.8106	5712	0.6550
Case-RP - After History-Matching	0.9715	9812	0.9869
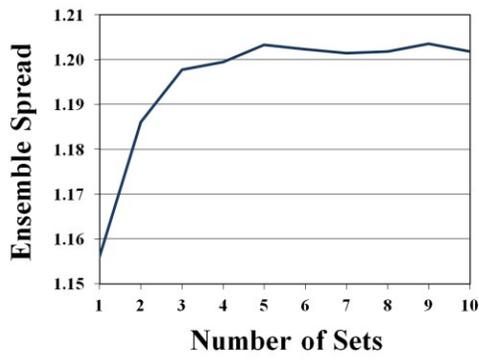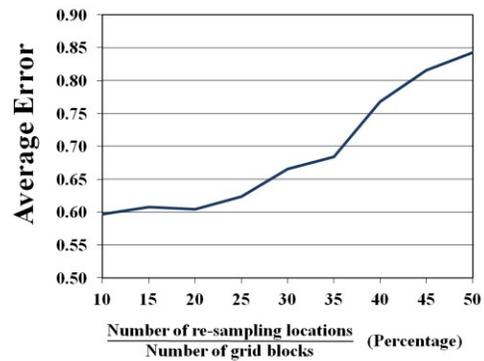
Figure 3-1 (a) Ensemble spread as a function of the number of re-sampling sets. (b) Weighted root mean square error of semi-variogram mismatch versus number of re-sampling locations. These plots are used to determine the optimum number of (a) re-sampling point sets and (b) total number of re-sampling points for Example A.
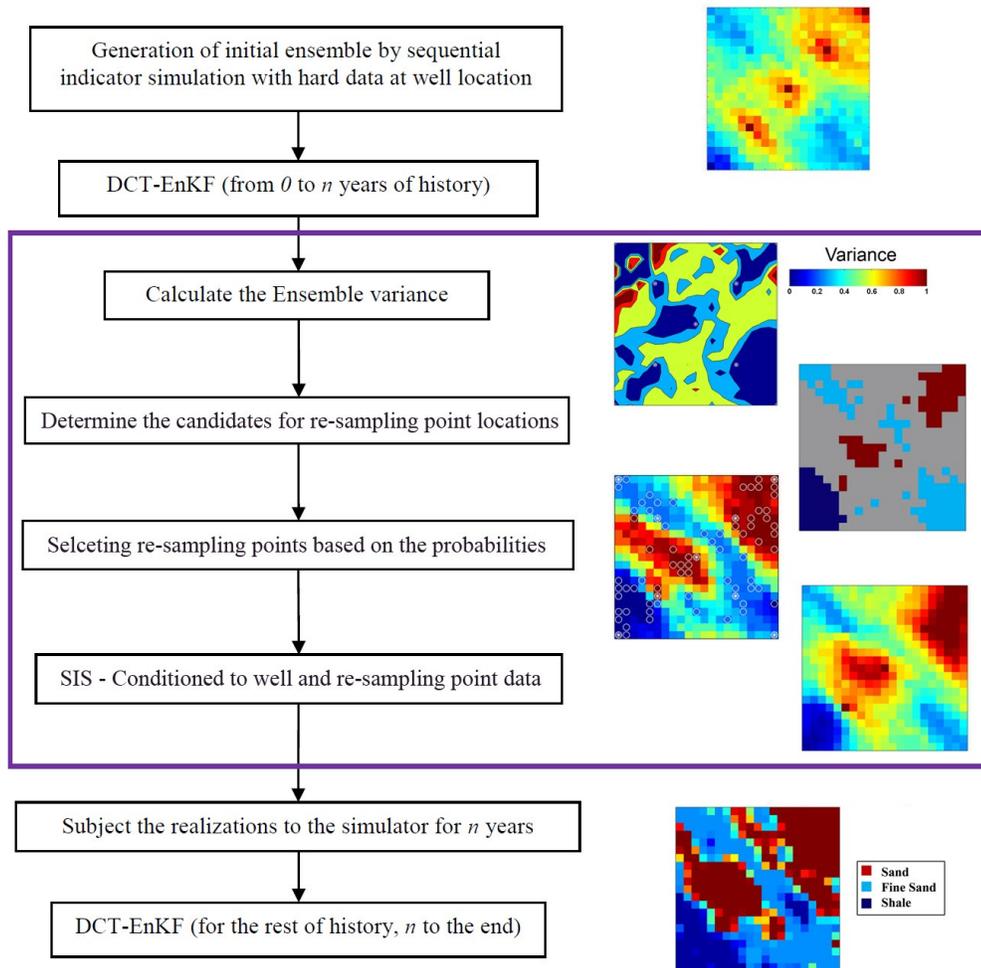
Figure 3-2 Diagram representing the workflow of Ensemble Kalman Filter with re-sampling and DCT parameterization.
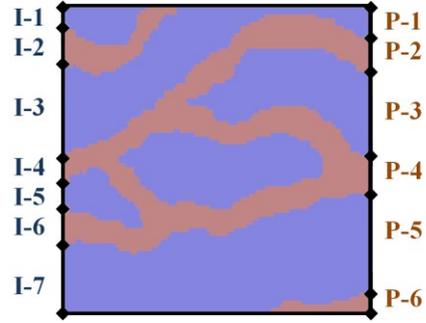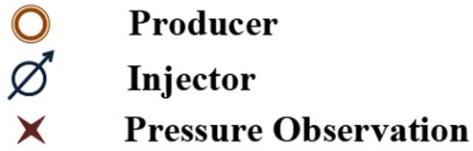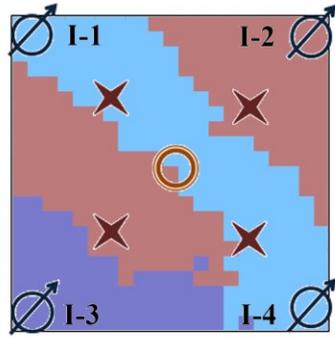
Figure 3-3 Well locations for both examples.
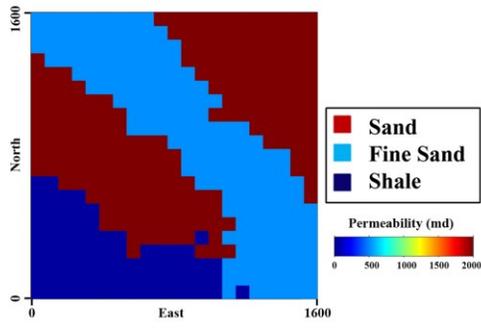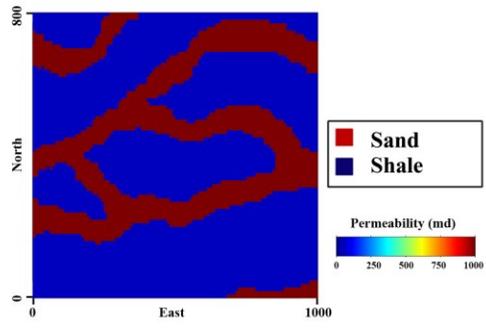
Figure 3-4 True Model for Example A (a) and Example B (b).
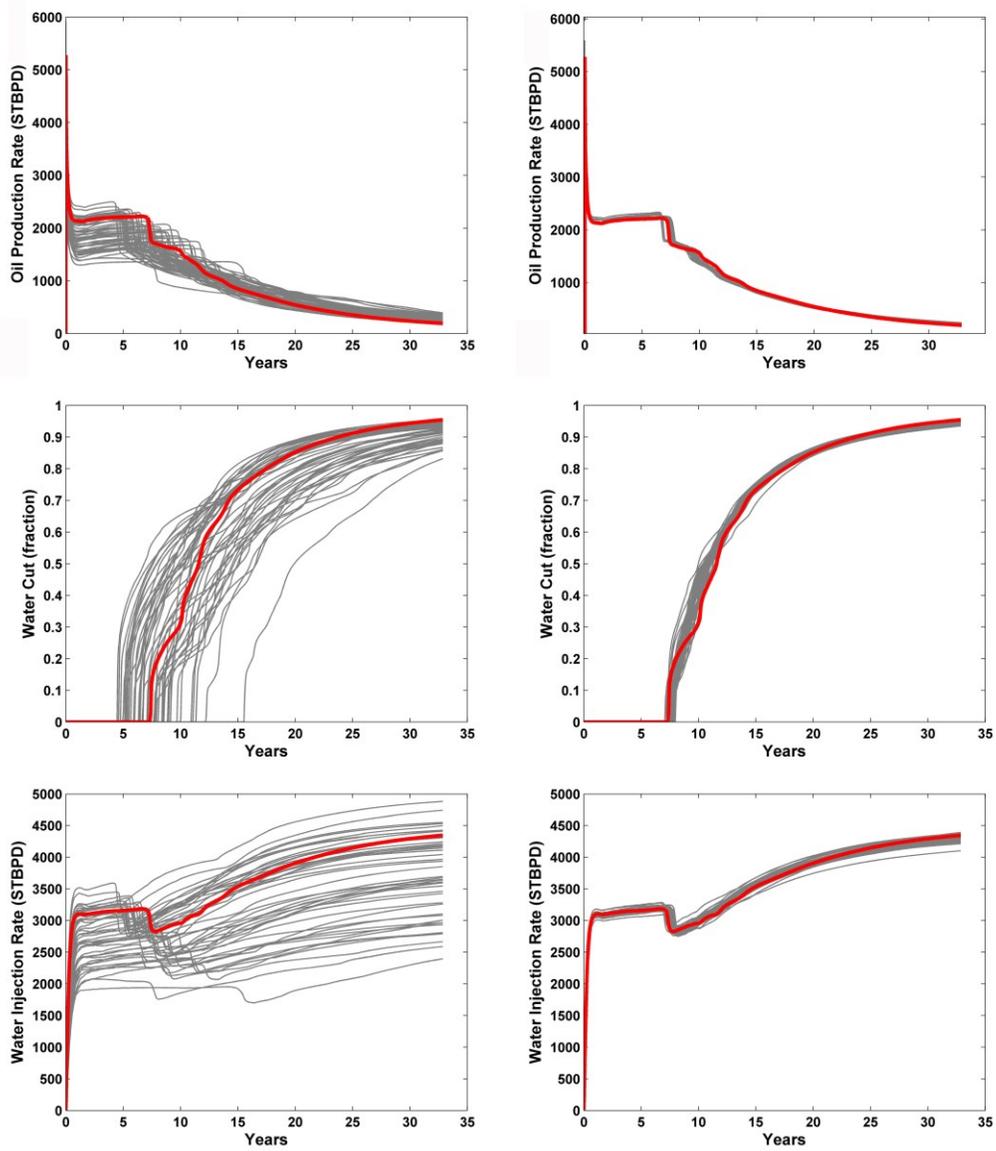
Figure 3-5 Oil production rate, water cut, and water injection rates before (left column) and after (right column) history match for Base Case. An acceptable match of observed data was achieved.

Base Case - Example A
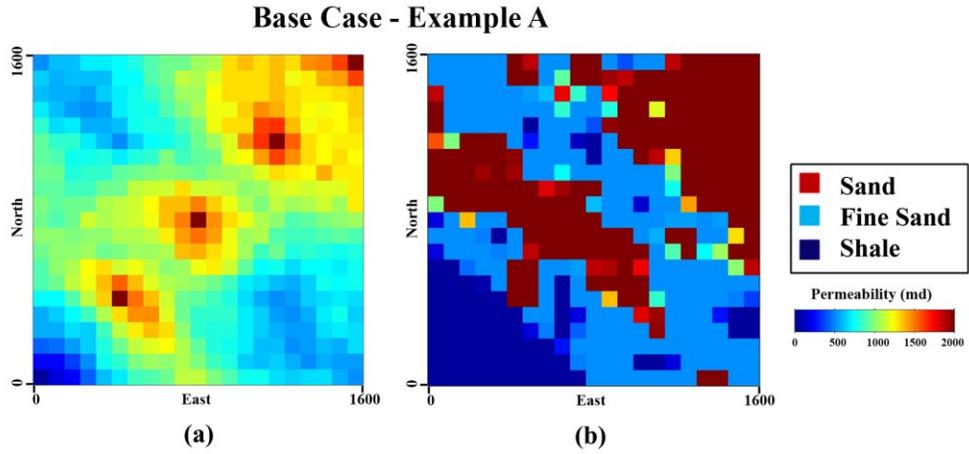
Figure 3-6 Average permeability map of the initial (a) and updated (b) ensemble (20 update steps, 10 years of production history).

Base Case - Example A
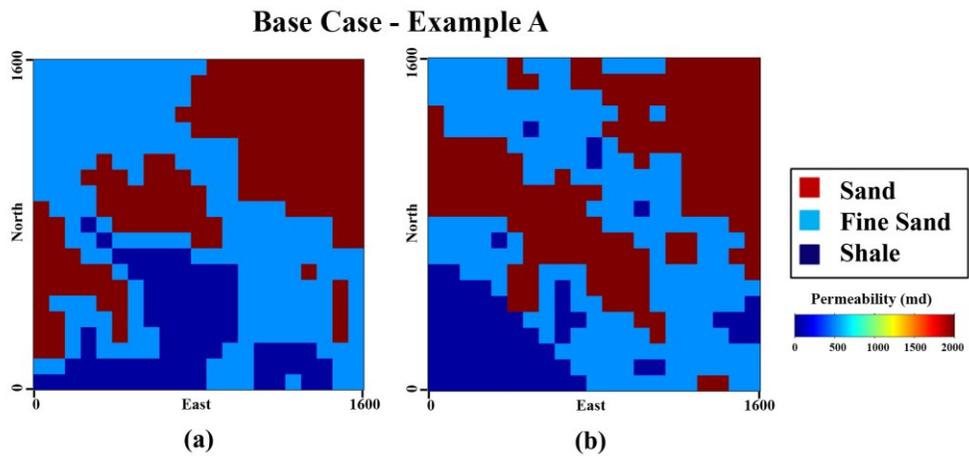
Figure 3-7 Permeability map of a randomly-selected ensemble member: initial (a) and final updated (b) model.
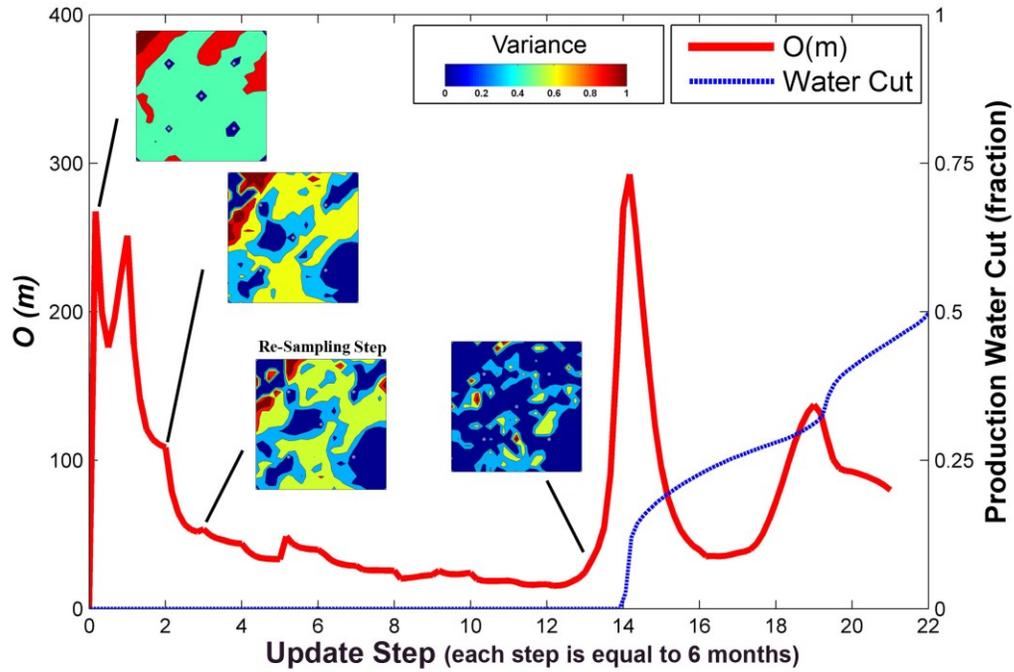
Figure 3-8 Objective function (Equation 3-19) and the actual production water cut history at successive assimilation steps. The inset figures show the ensemble variances of model parameters at various assimilation steps: 0, 2, 3, and 13 (t = 0, 12, 18, and 78 months).
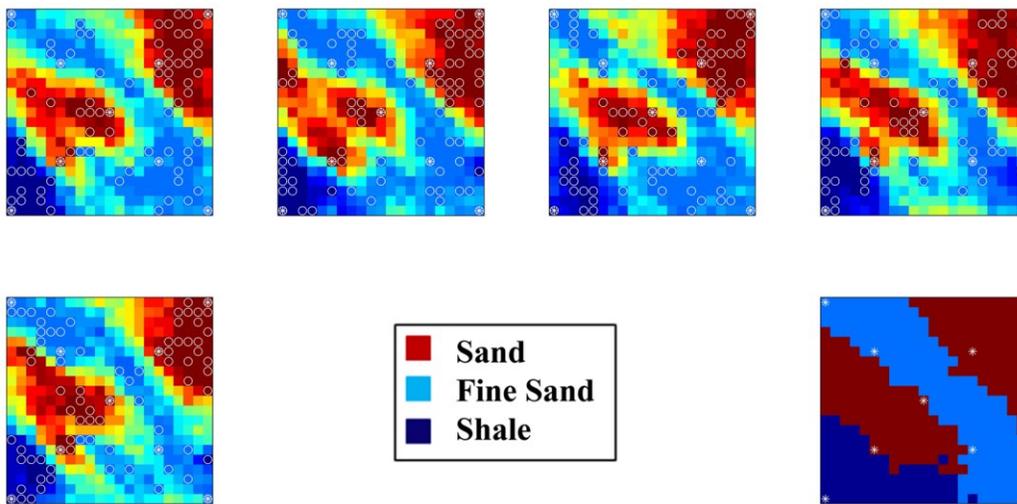
Figure 3-9 Comparison of the average permeability map of the five re-sampled ensembles to the true model as shown in Figure 3-4. Circles denote the locations of selected re-sampling points.
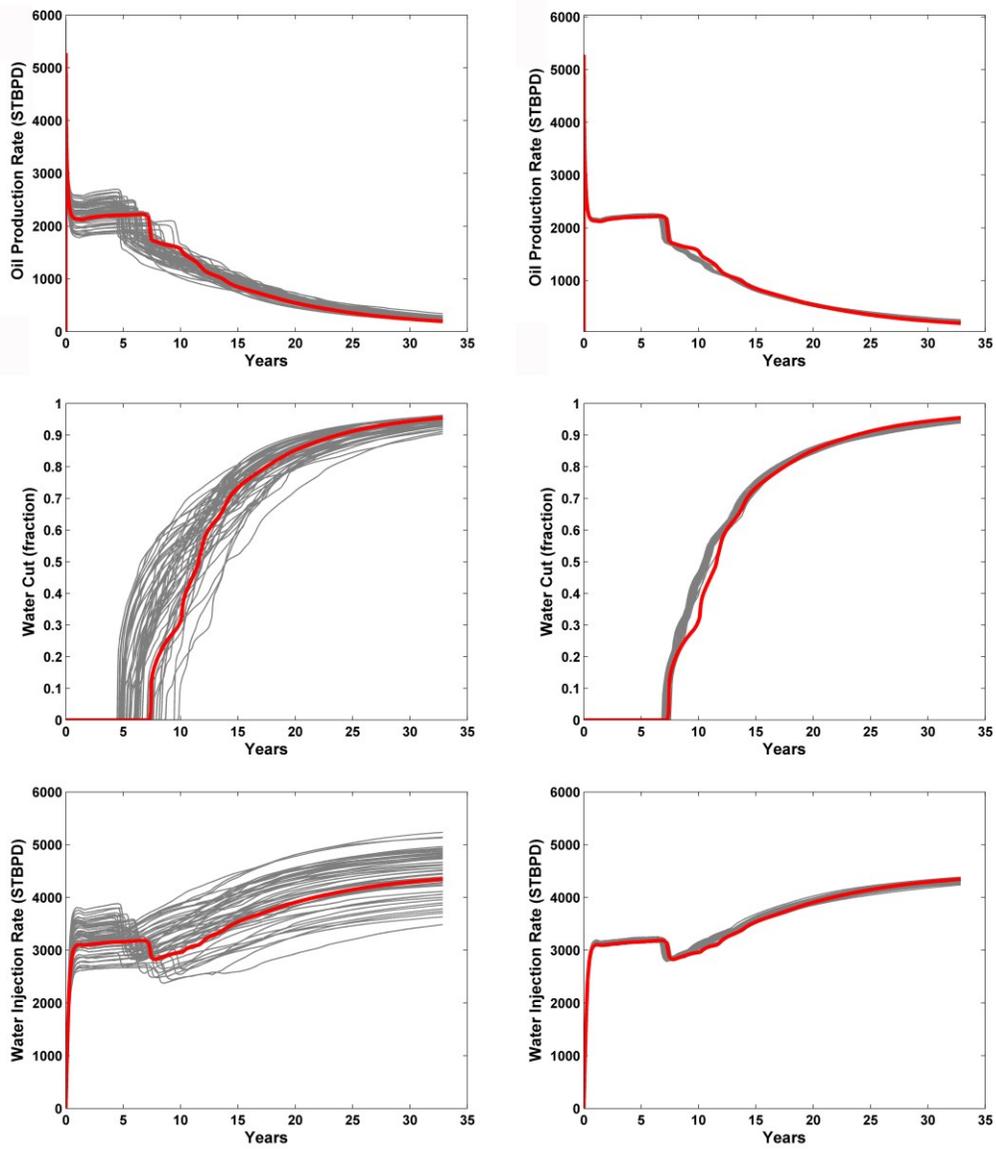
Figure 3-10 Oil production rate, water cut, and water injection rates of the re-sampled ensemble before (left column) and after (right column) resuming EnKF updating for Case-RP. Ensemble shown on the left was generated using probability weighted re-sampling (i.e., based on the results of Base Case after 3 update steps, 18 months of production history).
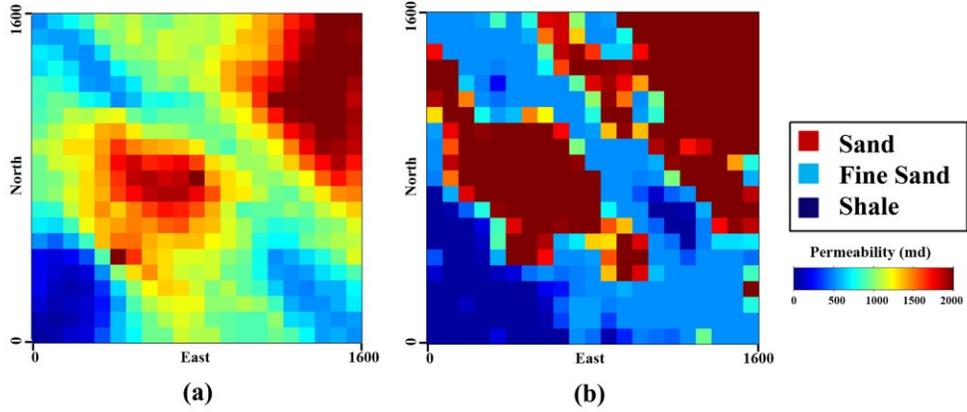
Figure 3-11 Average permeability map of the initial (a) and updated (b) ensemble (14 update steps, 7 years of remaining production history). Initial ensemble was generated by probability weighted re-sampling.
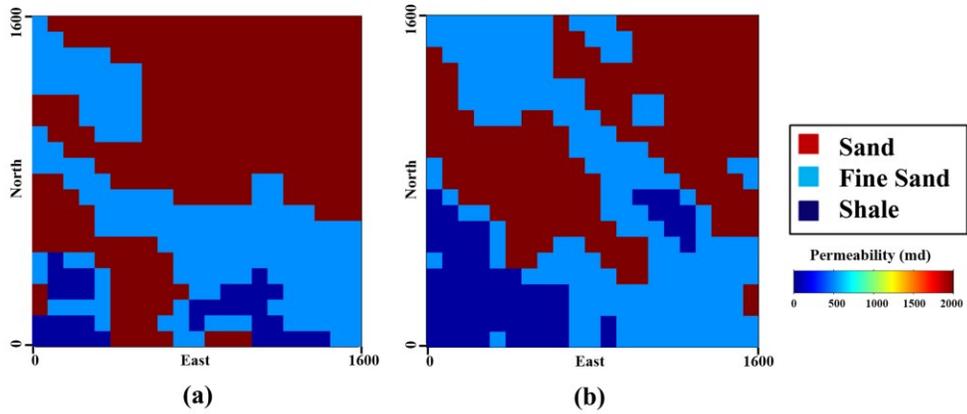
Figure 3-12 Permeability map of a randomly-selected ensemble member: initial (a) and the updated (b) model. Initial ensemble was generated using probability weighted re-sampling.

108
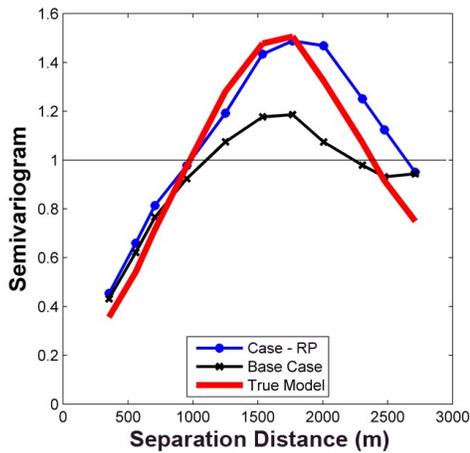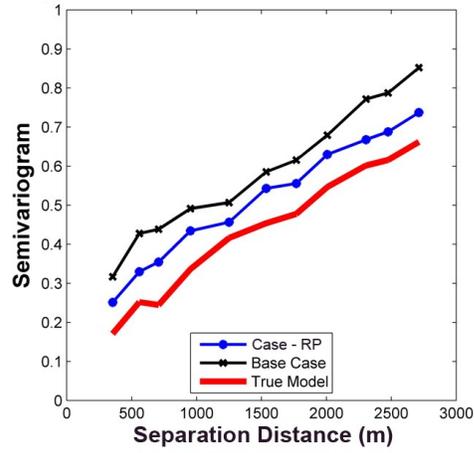
Figure 3-13 Comparison of experimental variogram values of a randomly-selected ensemble member for true model, base case, and re-sampling case, along the horizontal azimuths of 45 (a) and 135 (b) degrees.
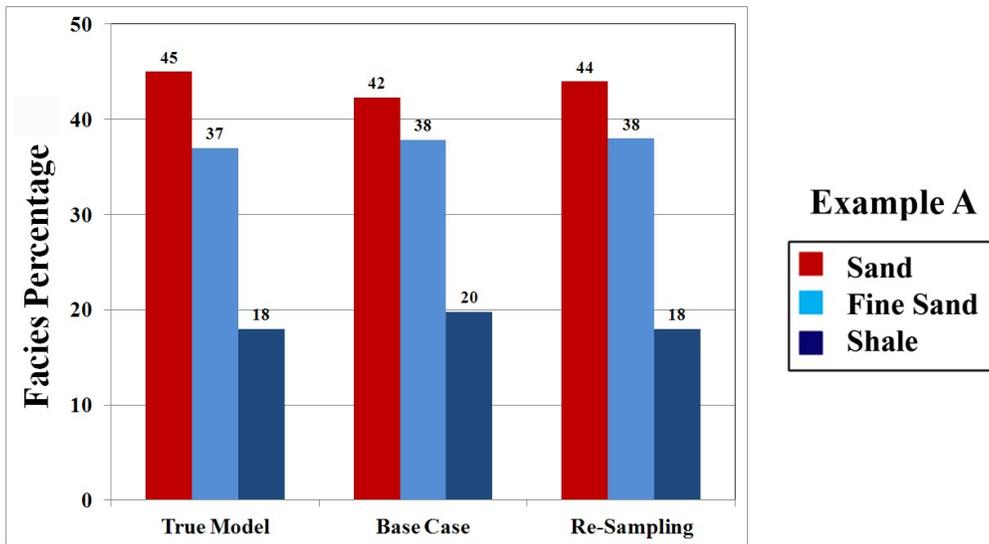
Figure 3-14 Comparison of experimental histogram of the average ensemble facies proportions in Example A.

109
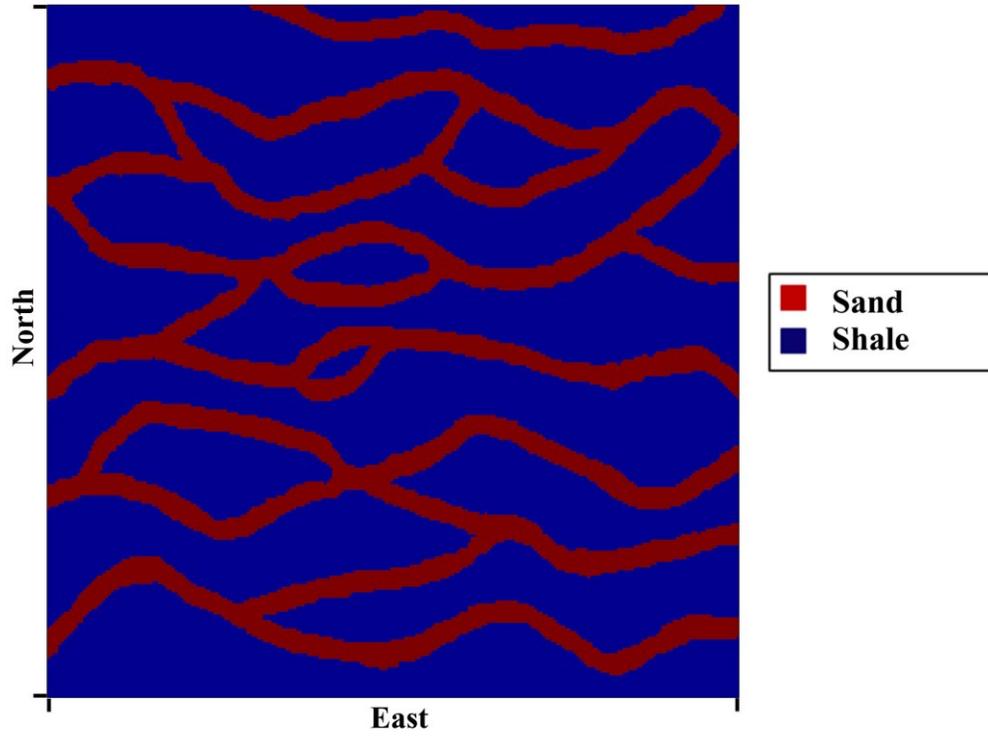
Figure 3-15 Training image for the channelized model in Example B.
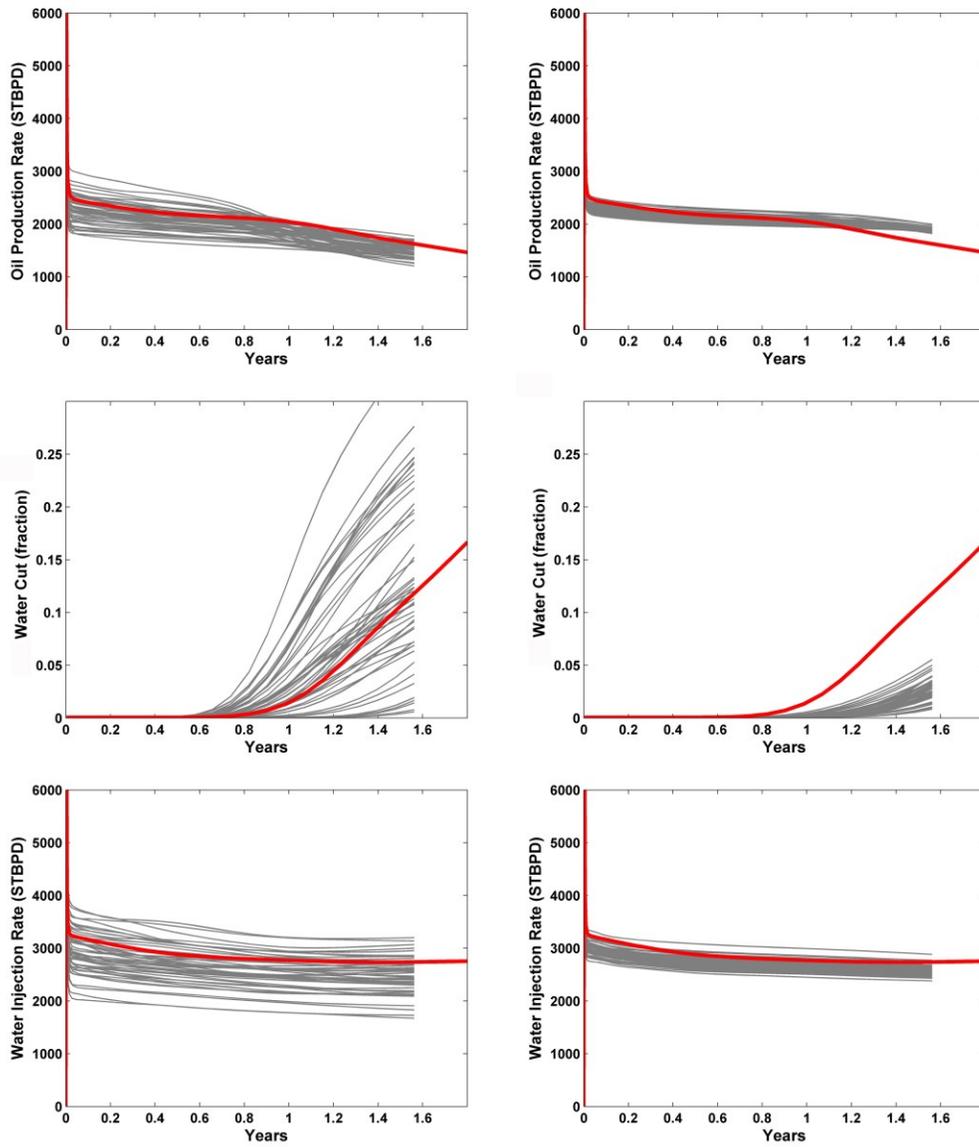
Figure 3-16 Oil production rates, water cut, and water injection rates before (left column) and after (right column) history match for Base Case. An acceptable match of observed data was achieved.
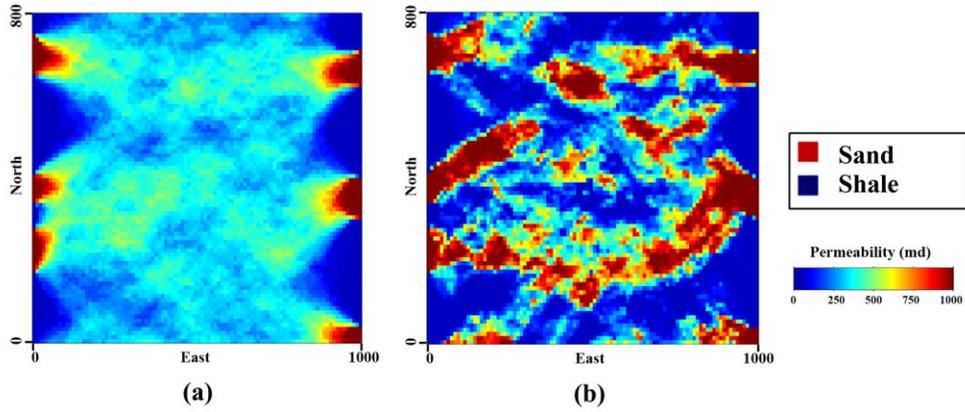
Base Case - Example B

Figure 3-17 Average permeability map of the initial (a) and updated (b) ensemble (9 update steps, 9 months of production history).

Base Case - Example B
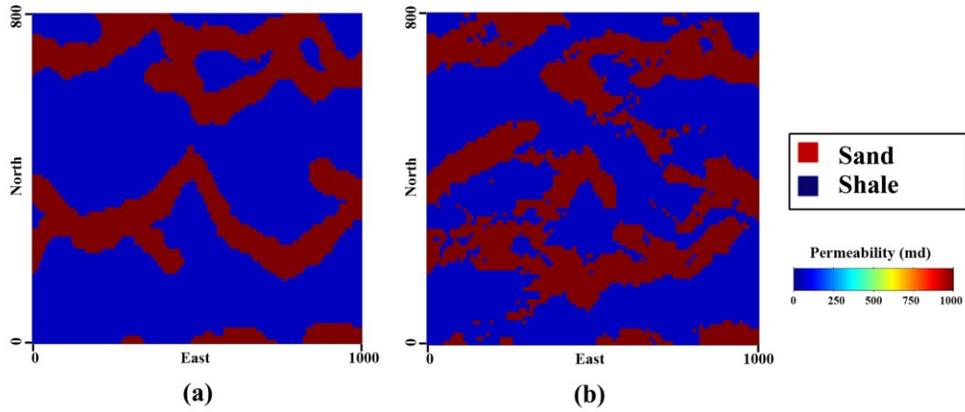
Figure 3-18 Permeability map of a randomly-selected ensemble member: initial (a) and final updated (b) model.
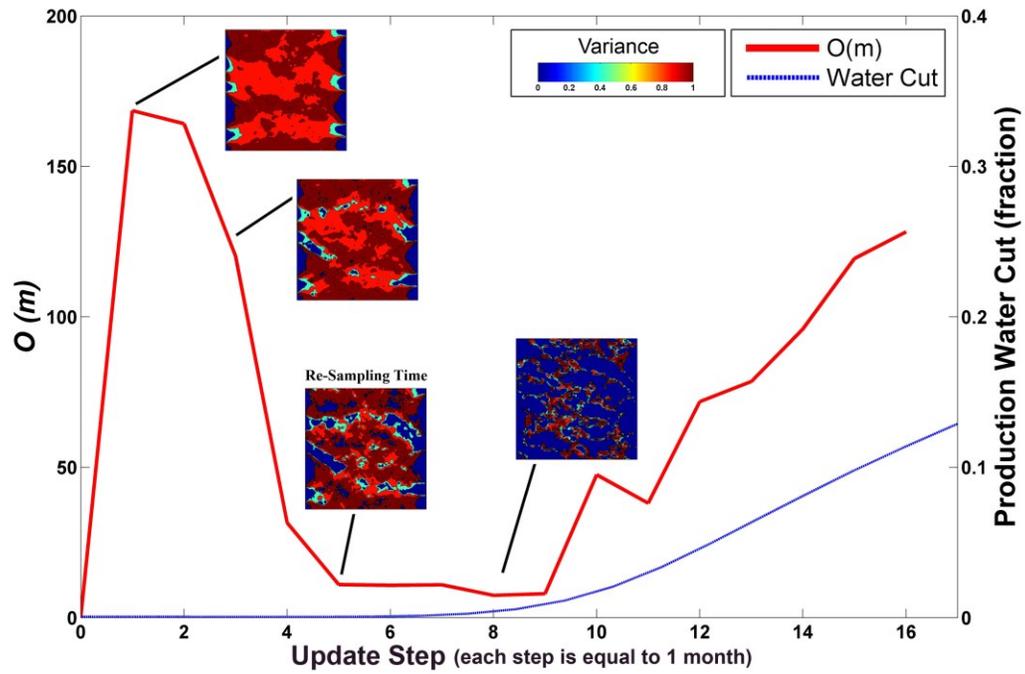
Figure 3-19 Objective function (Equation 3-19) and the actual production water cut history at successive assimilation steps. The inset figures show the ensemble variances of model parameters at various assimilation steps: 1, 3, 5, and 8.
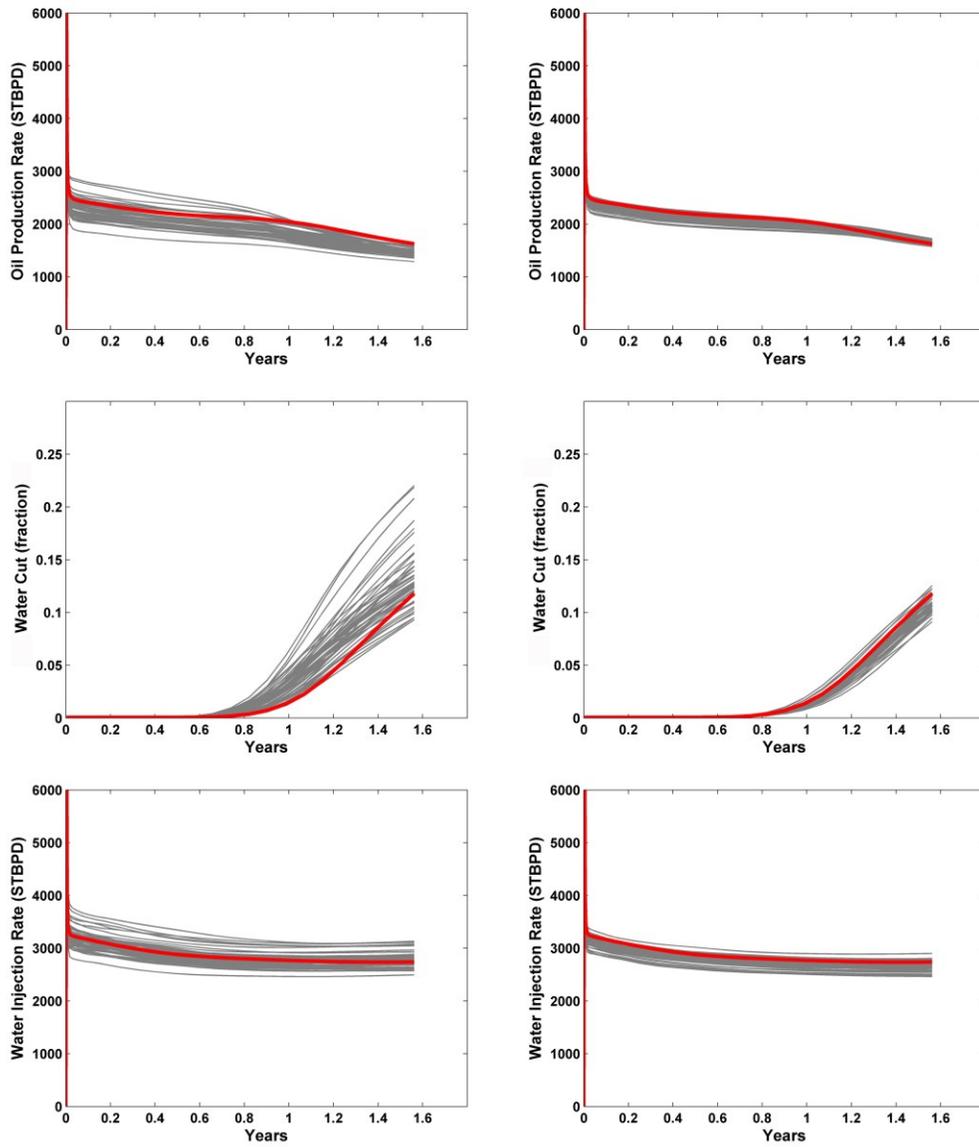
Figure 3-20 Oil production rates, water cut and water injection rates of the re-sampled ensemble before (left column) and after (right column) resuming EnKF updating for Case-RP. Ensemble shown on the left was generated by probability weighted re-sampling.
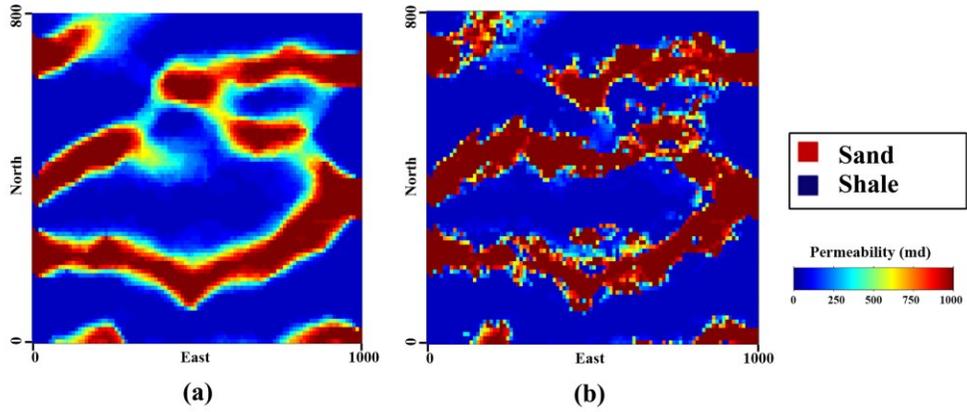
Figure 3-21 Average permeability map of the initial (a) and updated (b) ensemble (9 update steps, 9 months of remaining production history).
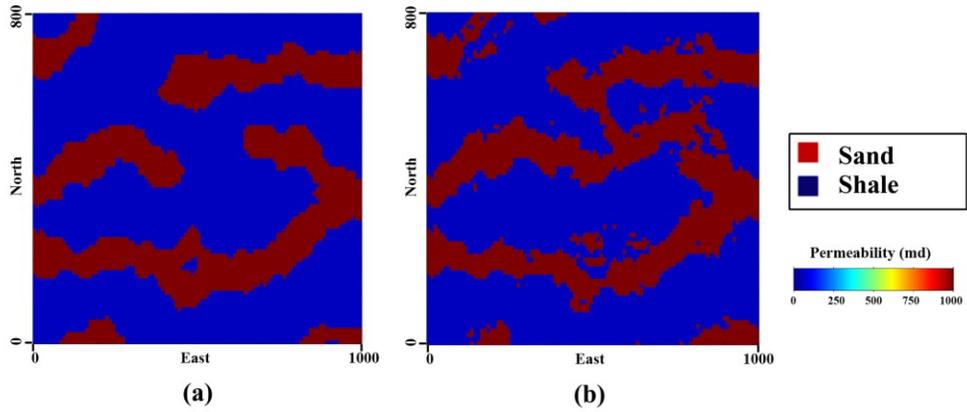
Figure 3-22 Permeability map of a randomly-selected ensemble member: initial (a) and the updated (b) model.

115
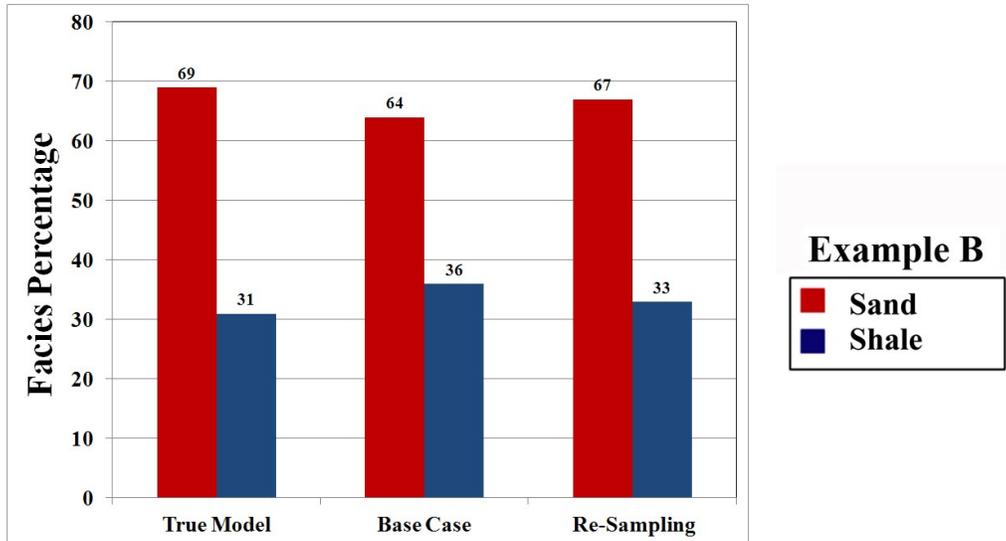
Figure 3-23 Comparison of experimental histogram of the average ensemble facies proportions in Example B.

Chapter 4: History Matching and Uncertainty Quantification of Discrete Fracture Network Models in Fractured Reservoirs[5]

4.1 Introduction

Fluid flow in fractured reservoirs mainly takes place through the network of interconnected fractures surrounding matrix blocks. Proper characterization of the fracture system is of outmost importance while making a robust model for simulation of the multiphase fluid flow in fractured reservoirs.

Continuum and discrete fracture modeling approaches are implemented for description of the fracture network. In the Continuous Fracture Modeling (CFM) method the geological and geomechanical factors that control fracturing of the formation are identified. These factors, known as fracture drivers, are related to fracture observation and indicators at the well locations. Once a relation is established, the fracture drivers are used to model the fractures throughout the reservoir and their equivalent porous medium parameters are further exported for use in numerical simulation. The Discrete Fracture Network (DFN) modeling approach focuses on modeling fractures explicitly. DFN models are characterized by properties such as length, orientation, intensity and transmissivity. Various data from different resources such as seismic reflection data, production tests, and well logs can be utilized to generate and validate the

---

DFN models and calibrate the flow properties of the fractures. However, due to computational complications and costs, DFN models consisting of thousands of fracture elements cannot be utilized to simulate multiphase flow at the full field scale (Dershowitz et al. 2004).

Developing a reliable and accurate simulation model is essential for predictions of hydrocarbon recovery and uncertainty assessment of fractured reservoirs. Dual-porosity modeling is widely adopted in fractured reservoir simulation (Warren and Root 1963, Kazemi et al. 1967). It conceptualizes the fracture-matrix system as dual continuous media in which the fluid flow takes place only through the high permeability fracture system and the matrix blocks act as the storage. An extension of dual porosity simulation is the dual-porosity dual-permeability approach, which allows fluids to flow through both the fracture network and between matrix blocks (Fung et al. 1991).

To construct a dual continuum reservoir simulation model, equivalent porous medium properties are assigned to each reservoir cell of both matrix and fracture continua. The determination of the equivalent parameters is an important step for simulating fluid flow in the fractured reservoirs. Different techniques have been introduced to upscale DFN to the equivalent dual porosity model. Various authors have extensively studied upscaling of fracture permeability data and the matrix fracture interaction. An analytical formation was developed by Oda (1985). Numerous authors have proposed other upscaling approaches (Bourbiaux et al. 1998, Karimi-Fard et al.2006, Bogdanov et al. 2007). Flow-based upscaling and local-global upscaling techniques have also been studied by a variety of

authors (Chen et al. 2003). Using upscaling techniques, individual fracture properties are upscaled to an equivalent permeability tensor and the sigma factor or shape parameter described in the dual porosity simulation model.

In addition to the static information (geologic and seismic data), history matching techniques are further implemented to integrate dynamic (production) data for enhanced reservoir characterization. Many techniques have been implemented for history matching of fractured reservoirs. The techniques focus on either directly updating dynamic model parameters, or updating fracture drivers and correlation parameters in CFM approach, or adjusting fracture parameters in DFN models.

Traditionally, fractured reservoirs are treated the same as conventional reservoirs, where field production performance is integrated to adjust dynamic model parameters. The history matching process updates parameters such as permeability, porosity and matrix fracture interaction coefficient. It ignores explicitly optimizing fracture parameters and updating complex geological features of fractured reservoirs.

Several authors have used the parameters of a volumetric grid, which is derived from continuous fracture models, as a tuning parameter. Ouenes et al. (1995) implemented a neural network to correlate geological information and well performance to the reservoir fracture intensity. Ouenes (2000) implemented a fuzzy neural network to evaluate the effect of different fracture drivers, such as structure, lithology and bed thickness on fractures and develop correlations between geological drivers and fracture intensity. The optimized fracture

intensity map is used for selecting potential infill drilling well locations with estimated ultimate recovery higher than certain economic limits. Sezuki et al. (2007) have used probability perturbation methods to adjust fracture intensity and large-scale fracture trend. They have calculated directional effective permeability of the single porosity reservoir simulation model from matrix permeability and fracture intensity. Cui and Kelkar (2005) have used a gradient simulator and adjoint method for conditioning the fracture intensity to the production data. They have used the fracture intensity map to estimate directional fracture permeability and matrix fracture interaction coefficient.

Few authors have focused on characterizing parameters of discrete fractures integrating dynamic production data. Gang and Kelkar (2008) have used the adjoint method to calibrate permeability of individual fractures as well as the capillary pressure curves. They have assumed that the relationship between the grid block effective permeability and fracture permeability is known. Hu and Jenni (2005) have used object-based Boolean simulation to define faults and fractures in the model. They have implemented the gradual deformation method to calibrate the object based model (location, shape and size of the objects) to dynamic data. De Lima et al. (2012) have used gradual deformation to characterize fault density, fault position and length of a fractal fault model.

This chapter attempts to address two particular issues with history matching of fractured reservoirs. Firstly, various fracture parameters, which are implemented in discrete fracture network modeling, are characterized. These parameters include global fracture parameters such as intensity and connectivity, as well as

single fracture parameters, such as orientation, size, shape, aperture, transmissivity, and storage. Prior to dynamic data integration, the probability distributions of the parameters are inferred from static geological evidence. The objective functions are further implemented in inverse modeling to minimize the observed data mismatch and obtain samples from the posterior probability density function (pdf). Current frameworks for optimization and history matching of fractured reservoirs lack the ability to estimate a large number of model variables. The applications mainly focus on updating few parameters.

Secondly, it is intended to generate a group of multiple history matched models, which are calibrated to both static and dynamic observations. A single history matched model might be useful; however, for quantifying uncertainty of the forecasts, multiple models with small dynamic data mismatch are necessary.

Oliver and Chen (2011) have reviewed various history matching techniques for quantifying uncertainty on production forecasts. Application of these techniques for fractured reservoir models and uncertainty assessment of fracture parameters is limited. Ghods and Zhang (2010, 2012) have used Ensemble Kalman filter (EnKF) for parameter estimation of dual porosity dual permeability reservoir models. Ping and Zhang (2013) have used EnKF combined with level set parameterization for estimating fracture distribution in a two dimensional synthetic reservoir model. Their principal objective was to capture the main features of the fracture distributions in the reference dynamic model. Cliffe et al. (2011) used Bayesian approximation to condition the transmissivity of a DFN model of a ground water flow problem to observed pressure data.

The Ensemble Kalman Filter (EnKF) is a Monte-Carlo technique that implements Bayesian updating scheme for data assimilation. It has been introduced to the petroleum engineering industry by Lorentzen et al. (2001) as a promising approach for solving high-dimensional history matching problems (Aanonsen et al. 2009). The EnKF implements an ensemble of model states and approximates the covariance matrices sequentially in time as new observations become available. The primary assumptions are linear system dynamics and that the model parameters follow a multivariate Gaussian probability distribution function (pdf). Applying the updating schemes iteratively during each assimilation step has been proposed to address the issues related to non-linearities (Chen and Oliver 2012, Li and Reynolds 2009). In the case of DFN characterization, the highly nonlinear relationship between the DFN model parameters and the simulation model production data implies that assimilating state variables (e.g. pressure data) is not practical and each proposed state requires a run of forward model after every update step to ensure consistency among state and model variables. The Gaussian assumption implies that the EnKF does not converge to the correct distribution, identified by geological studies, if the prior joint pdf has non-Gaussian contributions. The conventional EnKF method must be modified such that it can be applied for models whose parameters are not characterized by multivariate Gaussian distributions (Evensen 2003, Aanonsen et al. 2009). Nejadi et al. (2012) have implemented Normal Score transform as a parameterization technique to update decomposed permeability tensor parameters of a fractured reservoir along principal flow

directions. The technique has shown successful result in their studies, where the model parameters followed a bimodal distribution. This chapter proposes an innovative methodology based on ensemble-based history matching solutions to estimate probability distribution of DFN model parameters. This is achieved by updating various DFN model parameters using EnKF data assimilation technique. In this application, the uncertain parameters of global fracture distribution and connectivity as well as individual fracture properties are updated. To be more specific, the global parameters include global fracture intensity, number of fractures connected to the well and tributary drainage volume. The individual fracture parameters are local grid fracture intensity, fracture trend and plunge (orientation parameters).

The organization of this chapter is as follows: we begin our discussion with the formulation of DFN model parameter estimation integrating dynamic data. Next, we present the results obtained from the application of the proposed methodologies in a synthetic tight gas reservoir model. At the end, the conclusion and remarks drawn from the case study are discussed.

4.2 Methodology

4.2.1 Generating discrete fracture network models

Initially, the fracture model is discretized into volumetric cells. The cells are equivalent to the grid blocks of the dual porosity model of the reservoir simulator. The geocellular model is implemented in our methodology to integrate data from different resources (dynamic and static) and characterize

123

DFN models. Different local fracture parameters (such as intensity and orientation) are correlated to the cell attributes (seismic attributes, stress or strain tensor).

Generation of Initial DFN Realizations: The initial ensemble of DFN realizations is generated by defining specific global fracture intensity (p_{32G}), for different realizations. Generally P_{32} is the area of fractures per volume of rock mass. The fracture intensity at each and every grid block is defined by local fracture intensity (P_{32L}). If secondary data, such as seismic coherence are available, P_{32L} of the initial realizations is correlated to the secondary data, otherwise it is considered to be constant all over the model.

Generation of DFN Realizations Using Updated Parameters: Global fracture intensity (p_{32G}), local grid based fracture intensity (P_{32L}) and fracture orientation parameters (trend Θ and plunge Φ) are the parameters, which are updated during history matching. The optimized cell parameters are implemented in geocellular fracture generation algorithm as implemented in FracMan® software to generate an improved ensemble of DFN models. The geocellular algorithm uses the cell properties to generate fractures to a volumetric grid. Total number of fractures connected to the well (n_{fw}), tributary drainage volume (vol_{TDV}), and global fracture intensity of tributary drainage volume ($p_{32G-TDV}$) are other parameters that affect fluid flow and the pathways in the reservoir. These parameters are estimated during dynamic data integration and used as the conditioning data to generate new DFN realizations. Monte

124

Carlo simulation is implemented to generate DFN models that honor the target (updated) parameters.

The DFN models are conditioned to the fracture observations at the well locations. Image logs, core analysis and production logs provide valuable information, in relation to fracture parameters such as fracture opening and linear fracture intensity (p_{10}, which is the number of fractures divided by the length of scanline). After a DFN is generated, statistical analyses are performed to examine how closely the intersecting fractures match the observed p_{10} value at the well location. If necessary, Monte Carlo simulation is implemented to minimize the mismatch between the experimental p_{10} of the generated realizations and the target p_{10} value.

4.2.2 Upscaling

Equivalent grid cell permeability tensors are computed as parameters of a dual porosity reservoir simulator by implementing upscaling techniques. In our implementations we have used Oda's analytical approach, which is computationally faster than flow-based upscaling techniques.

4.2.3 Defining orientation parameters

Given a discrete fracture network model, upscaling is carried out to calculate the grid cell permeability tensors for reservoir simulation. The upscaling procedures will generally create full tensor permeability for each and every grid block of the

geocellular reservoir model. The permeability tensor is a symmetric and positive definite matrix. The permeability tensor, K, is in the following form:

$$K = \begin{bmatrix} k_{xx} & k_{xy} & k_{xz} \\ k_{yx} & k_{yy} & k_{yz} \\ k_{zx} & k_{zy} & k_{zz} \end{bmatrix}$$

In conventional homogenous reservoirs, the grid blocks of the reservoir simulation model are defined along three directions that transform the permeability tensor into a diagonal matrix. Specifically, a pressure drop along the coordinate directions would only yield to fluid flow in same path. In highly heterogeneous reservoirs, with lots of variations in rock properties, defining a local coordinate direction according to the fluid flow pathways is practically impossible. As a result, the off-diagonal permeability tensor elements of the grids in which the principal flow directions do not coincide with the grid coordinate system are not zero.

Eigen decomposition is implemented to factorize a matrix into canonical terms. The permeability tensor can be represented in terms of its eigenvalues and eigenvectors. The eigenvectors define the principal flow directions in the grid block and the eigenvalues describe how rapidly fluids can flow along the principal directions. Since the permeability tensor is a symmetric, positive definite matrix, its eigenvectors are orthogonal. The first two eigenvectors, corresponding to the two largest eigenvalues, describe the spatial orientation of a plane in which the ability of the rock to permit fluid flow is highest. In fractured

reservoir models, this plane, describes the average orientation of the fractures in a grid block. The third eigenvector is perpendicular to the prescribed plane and defines the average pole orientation of the fractures within the grid (Figure 4-1). In other words, the eigenvector corresponding to the smallest eigenvalue is the pole vector of the predominant fluid flow plane in the grid (Nejadi et al. 2012).

Consider the first eigenvector V_1, its components in the spherical space can be written as

$$V_1 = \begin{bmatrix} v_{x1} \\ v_{y1} \\ v_{z1} \end{bmatrix}, \quad \begin{aligned} r_1 &= 1 \\ \theta_1 &= \tan^{-1}\left(v_{y_1}/v_{x_1}\right) \\ \varphi_1 &= \cos^{-1}\left(v_{z_1}/r_1\right) \end{aligned} \qquad \text{4-1}$$

Consider the second eigenvector V_2, its components in spherical space can be written as

$$V_2 = \begin{bmatrix} v_{x2} \\ v_{y2} \\ v_{z2} \end{bmatrix}, \quad \begin{aligned} r_2 &= 1 \\ \theta_2 &= \tan^{-1}\left(v_{y_2}/v_{x_2}\right) \\ \varphi_2 &= \cos^{-1}\left(v_{z_2}/r_2\right) \end{aligned} \qquad \text{4-2}$$

Similarly, the angles of the third eigenvector V_3 in spherical space can be written as

$$V_3 = \begin{bmatrix} v_{x3} \\ v_{y3} \\ v_{z3} \end{bmatrix},$$
$$r_3 = 1$$
$$\theta_3 = \tan^{-1}\left(v_{x_3} / v_{y_3}\right)$$
$$\varphi_3 = \cos^{-1}\left(v_{z_3} / r_3\right),$$
4-3

where θ_3 and φ_3 are the trend and plunge of pole vector, which is perpendicular to the predominant fluid flow plane in each grid.

The permeability tensors of all grid blocks in the upscaled model are decomposed into its principal directions, above formulation is implemented to calculate the trend and plunge of the pole vector. The trend and plunge values are updated by means of optimization algorithms and are further used as soft data for an improved estimation of fracture orientations in DFN models

$$\Theta = \left[\theta_3^1, \theta_3^2, \theta_3^3, ..., \theta_3^{n_b}\right]^T,$$
4-4

$$\Phi = \left[\varphi_3^1, \varphi_3^2, \varphi_3^3, ..., \varphi_3^{n_b}\right]^T,$$
4-5

where n_b is the total number of grid blocks in the geocellular model and T denotes the matrix transpose.

4.2.4 Forecast model

In a fractured reservoir, fluids exist in two connected systems: the matrix and the fractures. Typically fractures provide high conductivity conduits and the matrix provides the fluid storage. A dual porosity/single permeability model may characterize a fractured reservoir with relatively low matrix permeability, but

high porosity. In this study, the commercial reservoir simulation software Eclipse (2011) was used as the forecast model.

The forecast model is separately applied to each ensemble member. The multiphase fluid flow in petroleum reservoirs is described by material balance, momentum balance, phase behavior descriptions and numerous auxiliary equations. This system of non-linear differential equations is solved numerically using methods such as finite difference

$$\begin{bmatrix} u_k^j \\ d_k^j \end{bmatrix} = G \begin{bmatrix} m_k^j \\ u_0^j \end{bmatrix} \ , j = 1, ..., n_e \ ,$$

4-6

where u_k^j denotes the estimated state variables (pressure and saturation) and d_k^j is the production data predictions, after running the dual porosity dual permeability reservoir simulator (forward model, G) from the initial conditions at time 0 to time step k. Individual realization is denoted by j, and n_e is the total number of ensemble members.

4.2.5 Ranking the realizations

After applying the forecast model to the upscaled DFN models, the simulations results are compared with the actual production data. The objective function is defined as the dynamic data mismatch or the root mean square error (RMSE)

$$O(m)_k^j = [d_k^j - d_{obs}]^T C_D^{-1} [d_k^j - d_{obs}] \ ,$$

4-7

where C_D is the error covariance matrix. The prescribed user defined objective function is implemented in the concept of importance weights in particle filter

technique, where ensemble members (particles) are weighted on the basis of likelihood function. Emerick and Reynolds (2012) suggested a weighting function that is proportional to $\exp(O(m))$. Nejadi et al. (2014) have also successfully adopted an analogous probability weighting scheme in the EnKF framework to maintain the reference statistics of the model parameters and to improve ensemble variability. A weighting scheme similar to that of Emerick and Reynolds (2012) is implemented to calculate the weight of each ensemble member. The approach is analogous to the use of importance weights in particle filtering techniques, where ensemble members (particles) are weighted on the basis of likelihood function. Ensemble members are promoted based on their likelihood and sampled members are used as the basis for optimization of fracture model parameters.

4.2.6 Analysis step – Ensemble Kalman filter

The analysis step is carried out using the observation data (d_k) for all ensemble members

$$y_k^{a,j} = y_k^{p,j} + KG_k \left(d_{ik} - Hy_k^{p,j} \right),$$

4-8

where $y_k^{a,j}$ represents the analysed state vector, $y_k^{p,j}$ is the posterior state vector, KG_k is the Kalman Gain at k^{th} step, and H is the observation operator, which represents the relationship between the state vector and the observation vector

$$H = \begin{bmatrix} 0 & | & \mathbf{I} \end{bmatrix}.$$
$$\hspace{10cm} \text{4-9}$$

d_{ik} is the observation data at the kth step ($d_{obs,k}$) plus observation noises (ε_k) for ensemble member i

$$d_{ik} = d_{obs,k} + \varepsilon_k.$$
$$\hspace{10cm} \text{4-10}$$

Kalman gain KG_k is defined as

$$KG_k = C_{y_k^{p,j}} H^T \left(HC_{y_k^{p,j}} H^T + C_{d_k} \right)^{-1},$$
$$\hspace{10cm} \text{4-11}$$

where $C_{y_k^{p,j}}$ is the state cross covariance matrix, and C_{d_k} is the error covariance matrix.

The cross covariance matrix is approximated as

$$C_{y_k^{p,j}} \approx \frac{1}{N_e - 1} \left\{ \left(y_k^p - \overline{y_k^p} \right) \left(y_k^p - \overline{y_k^p} \right)^T \right\},$$
$$\hspace{10cm} \text{4-12}$$

where T indicates matrix transpose, and $\overline{y_k^p}$ is the average of all posterior state vectors.

In this study, the state vector of the ensemble Kalman filter formulation (y_k^j) consists of the DFN model parameters (m_k^j) and the production data predictions (d_k^j), which are the simulation outputs of the dynamic model.

$$y_k^j = \begin{Bmatrix} m_k^j \\ d_k^j \end{Bmatrix}.$$

4-13

The following parameters are included in the state vector as the model parameters (m_k^j) and updated using EnKF algorithm:

- Global fracture intensity of the model (p_{32G}, which is scalar parameter corresponding to each realization)

- Local grid cell fracture intensity (P_{32L}, which is $n_b \times 1$ column vector corresponding to each realization. n_b is the total number of grid blocks in the geocellular model)

- Fracture trend (Θ, which is $n_b \times 1$ column vector corresponding to each realization)

- Fracture plunge (Φ, which is $n_b \times 1$ column vector corresponding to each realization)

- Total number of fractures connected to the well (n_{fw}, which is scalar parameter corresponding to each realization)

- Tributary drainage volume vol_{TDV} and fracture intensity of the tributary drainage volume ($p_{32G-TDV}$, which is scalar parameter corresponding to each realization)

$$m_k^j = [p_{32G}, P_{32L}{}^T, \Theta^T, \Phi^T, n_{fw}, vol_{TDV}, p_{32g-TDV}]^T{}_k^j.$$

4-14

4.3 Procedure

The detail procedure to update DFN models and their upscaled flow simulation models is described below:

a) Initially the following parameters are defined to generated the DFN realizations:

1- p_{32G} is derived from a uniform distribution. The true expected value of the total number of fractures is believed to lie between the lower and upper limit of the distribution.

2- P_{32L} is assumed to be constant for all grid cells.

3- Initial values for orientation parameters (Θ, Φ) are obtained from fracture observations at the well locations and other geological information.

4- n_{fw}, vol_{TDV} and $p_{32G-TDV}$ are unknown and DFN realizations are only conditioned to observed p_{10} value at the well locations.

b) Generate ensemble of DFN realizations;

1- p_{32G}, P_{32L}, Θ and Φ are implemented in geocellular fracture generation algorithm.

2- Using Monte Carlo simulation, the realizations are conditioned to the observed p_{10} value; at subsequent assimilation steps, these realizations are also conditioned to the updated n_{fw} and $p_{32G-TDV}$.

133

c) Perform statistical analysis to calculate the following statistics for the new DFN realizations:

    1- P_{32L}

    2- n_{fw}

    3- vol_{TDV}

    4- $p_{32G-TDV}$

d) Upscale DFN realizations.

e) Estimate Θ and Φ from Equation 4-4 and Equation 4-5.

f) Apply the forward model (G) to the realizations (from $t_0=0$ to t_k, k is the number of update step).

g) Calculate the objective function (Equation 4-7) and rank the realizations.

h) Apply EnKF to the high rank realizations and estimate updated model parameters.

i) Repeat step b to h for the entire production history.

The aforementioned procedure is schematically presented in Figure 4-2.

4.4 Case Study

The proposed methodology has been implemented for history matching of a synthetic tight gas fractured reservoir model. Three-stage hydraulic fracturing operation is used for well treatment and improving the productivity. Parameters of the hydraulic fracture planes are assumed to be known. The objective of this case study is to predict the spatial distribution and orientation of the natural fractures in the reservoir.

134

4.4.1 Reference model description

The dimensions of the model are 3000 m × 3000 m × 30 m. One horizontal well (700 m long) is located in the model. The well and hydraulic fracture locations are shown in Figure 4-3.

One set of natural fractures is present in the reservoir. The global model fracture intensity (p_{32G}) of the reference model is 0.03 m^{-1}. The fracture orientation parameters of the fracture pole vector trend and plunge are 85° and 15° respectively. The orientation distribution of the pole vector is defined using Fisher (1953) distribution; the Fisher dispersion parameter is set to 15. The fracture size distributions are derived from uniform probability density functions; the minimum and maximum fracture size is 66 m and 135 m respectively, while the aspect ratio is ten. The spatial fracture distribution of the reference model is depicted in Figure 4-4. Table 4-1 summarizes the parameters of the reference model.

The numerical model for the case study consists of a 100 × 100 × 10 grid in x-y-z directions. A dual porosity model is considered for reservoir simulation. The matrix properties are assumed to be constant throughout the model. Matrix porosity and permeability are constant and defined as 5% and 0.01md respectively. The fracture system properties are derived from upscaling the DFN model, implementing the Oda (1985) analytical technique. Fracture permeability (in y direction, k_{yy}) for the dual porosity simulation model is depicted in Figure 4-5.

The production schedule is defined for 28 months, and the complete production history is used for history matching and calculating the RMSE. The producer is operated with a maximum gas production rate of 141.6 standard cubic meters (SCM) per day and a minimum bottom hole pressure limit of 3.45e+6 Pa. The synthetic field-observed data set, which consists of gas production rate and bottom-hole pressure, is obtained by subjecting the reference model to a reservoir simulator. The Eclipse black-oil simulator was used as the reservoir simulator (forward model). Table 4-1 summarizes the parameters of the discrete fracture network model and the dynamic dual porosity model.

4.4.2 Implementation of history matching algorithm

To begin, 150 independent realizations, all equally probable, are generated. The individual DFN realizations are created using geocellular fracture generation algorithm. The global fracture intensity (p_{32G}) of the DFN models is derived from a uniform distribution. The lower and upper bound of the distribution are 0.005 and 0.09 (m^{-1}), respectively. The local fracture intensity (P_{32L}) is considered to be constant all over the reservoir model.

After generating the DFN models, local fracture intensity (P_{32L}) of the models is calculated. P_{32L} map of four randomly selected realizations are shown in Figure 4-6. Statistical analysis is implemented to calculate total number of fractures connected to the well and the tributary drainage volume. Figure 4-7 shows the histogram of total number of fractures connected to the well (n_{fw}), tributary drainage volume (vol_{TDV}), and the global fracture intensity within

tributary drainage volume ($p_{32G-TDV}$), for the initial ensemble members. Figure 4-8 shows the equal area (Schmidt), lower hemisphere stereonet of fracture pole orientations of a randomly-selected initial ensemble model. Probability distribution function of the fracture orientation follows Fisher (1953) distribution.

In the next step, the DFN models are upscaled. The upscaled models are subjected to the forward model (dual porosity, Eclipse reservoir simulator) for the first production period ($t_0 = 0$ to $t_1 = 4 months$). Figure 4-9 shows the cumulative gas production and well bottom-hole pressure of the initial ensemble compared to the true production values.

Equation 4-7 is used to compare the production performance of individual realizations with the true production history. The calculated objective function value, which is equivalent to the root mean square error, is implemented to assign a weight to the realizations. The weighting function is proportional to exp(-*O(m)*), where *O(m)* is a user-defined objective function. The weighting factors are used to sample from the realizations. Probability sampling is implemented to select members for the subsequent update steps.

History matching is performed for 7 steps. Update steps are every 4 months. The reservoir simulation is applied from $t_0 = 0$ to t_k where k is equal to 4 months for the first step and 8, 12… and 28 months for the subsequent update steps. The updated local fracture intensity maps (P_{32L}) of four randomly selected realizations are shown in Figure 4-10.

Figure 4-11 shows the histogram of total number of fractures connected to the well, tributary drainage volume, and the global fracture intensity within tributary drainage volume, for the updated realizations. Figure 4-12 shows the equal area (Schmidt), lower hemisphere stereonet of fracture pole orientations of a randomly-selected updated ensemble model. Figure 4-13 shows the cumulative gas production and well bottom-hole pressure of the updated models compared to the true production values.

4.5 Results and Discussion

Root mean square data were used to measure the quality of the observed data match and to assess the performance of the methodologies (compare Figure 4-9 and Figure 4-13) Furthermore, the parameters of the updated discrete fracture network models are in agreement with the geological parameters of the formation (the true model) and realistic fracture parameters distributions are maintained in the updated ensemble (Figure 4-11). In the case study, it is illustrated that several equally probable dynamic models match the production history. These realizations and their equivalent DFN models are implemented for reservoir production forecasting and uncertainty assessment of the fracture parameters.

Figure 4-7 and Figure 4-11 depict the histogram of the number of fractures connected to the well in the initial and updated ensemble. These fractures highly affect the average reservoir permeability in the region around the well and in the tributary drainage volume. The updated models clearly reveal the expected number of fractures which are connected to the well and the tributary drainage

138

volume for our case study. By analogy, the fracture intensity maps of the updated ensemble members (Figure 4-10) should be compared against the initial ensemble members (Figure 4-6). Prior to dynamic data integration, few static evidences are typically available to approximate the fracture intensity and its relation to geological parameters. The updated P_{32L} maps demonstrate the expected bounds and spatial distribution of the fracture intensity in the case study. These data are potentially implemented to update the relation between geological drivers and fracture intensity in the region far from the wells. Table 2 summarizes the statistics of some important DFN model parameters, which were updated during the history matching.

4.6 Conclusion

The proposed new methodology, integrates static data and dynamic observations for characterization and history matching of fractured reservoirs. It implements the root mean square error of production data match to update discrete fracture network model parameters and propose multiple equally probable DFN models and their equivalent continuum dual-porosity models. The updated models match the observed dynamic data and reproduce geologically realistic fracture parameters.

Fracture parameters having the highest uncertainty and impact on the production performance are selected as optimization parameters. Subsequent to each optimization step, a new ensemble of DFN models are generated using updated parameters. This aids the algorithm to maintain the diversity among ensemble

members and avoid underestimation of the uncertainty in the posterior probability distribution. The technique presents a promising potential in characterization of fractured reservoir model parameters using diverse data sources.

Table 4-1 Parameters of the discrete fracture network model and the dynamic dual porosity simulation model.

Discrete fracture network parameters:	
Fracture aperture	1e-6 m
Fracture transmissivity	1000 m²/hr
Fracture storativity	1 e-6 1/m
Dynamic model specifications:	
Model dimensions	100×100×10 grids in x, y and z directions
Grid dimensions	30 × 30 × 3 m
Reservoir depth	1220 m
Matrix porosity	5% constant
Matrix permeability	0.01md constant
Initial reservoir pressure	1.37895e+7 Pa
Well constraints:	
Maximum gas productions rate	141.584 SCM/day
Minimum bottom hole pressure	3.44738e+6 Pa
Total generated history	28 months

Table 4-2 Mean and standard deviation (STD) of different discrete fracture network model parameters for the initial ensemble compared with the updated realizations.

Parameter	Initial Models	Updated Models	True
$p_{32\,G}$	Mean: 0.0360 STD: 0.0143	Mean: 0.0292 STD: 0.0032	0.0293
n_{fw}	Mean: 296 STD: 388	Mean: 89 STD: 42	47
vol_{TDV}	Mean: 5.16e+7 STD: 5.77e+7	Mean: 1.79e+7 STD: 0.79e+7	1.10e+7
$p_{32\,G\text{-}TDV}$	Mean: 0.0375 STD: 0.0085	Mean: 0.0392 STD: 0.0062	0.0359
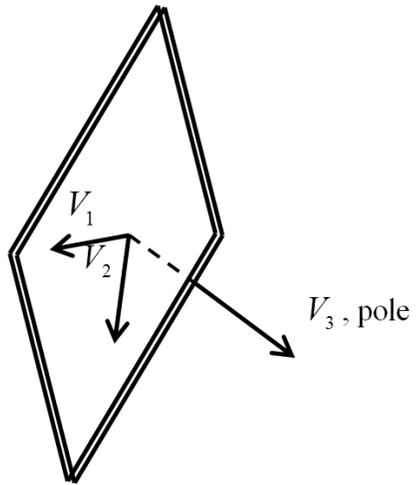
Figure 4-1 Schematic diagram showing fracture plane, V_1, V_2, and V_3 which is the fracture pole vector.
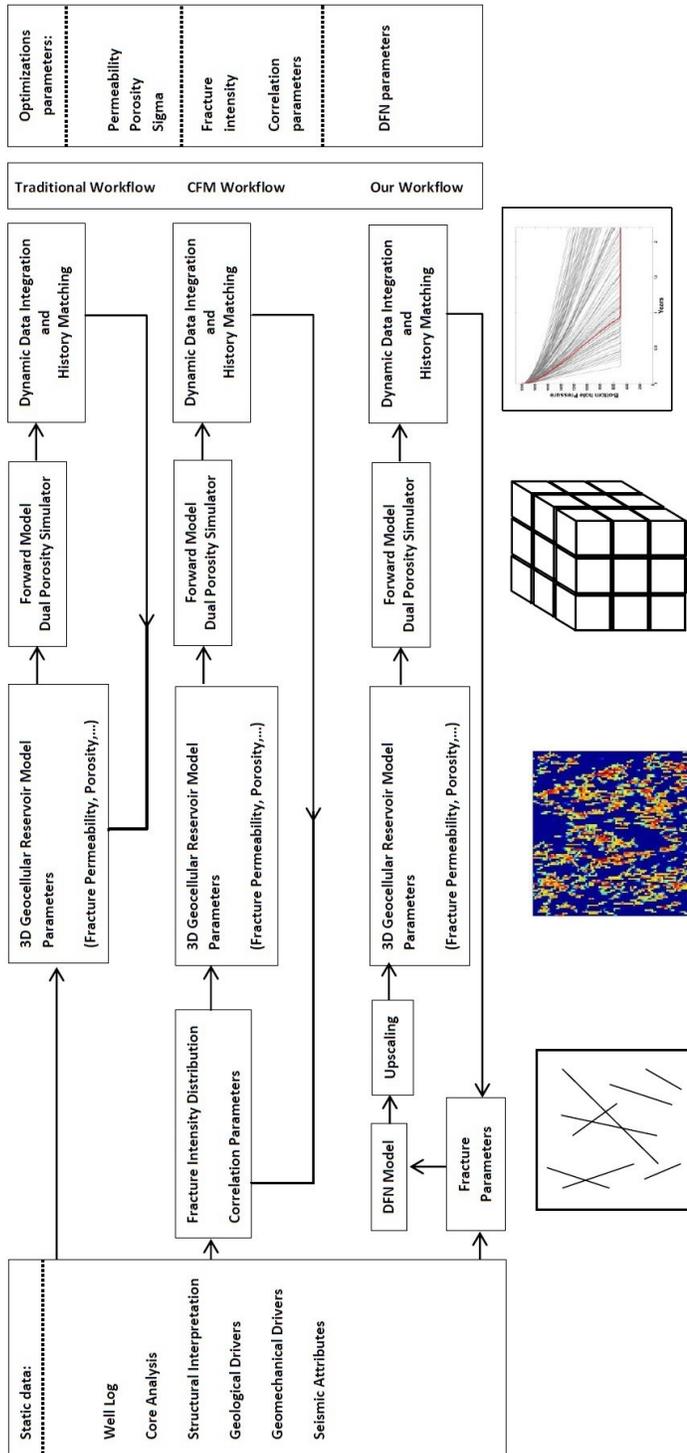
Figure 4-2 Diagram representing various steps for characterization of fractured reservoirs.
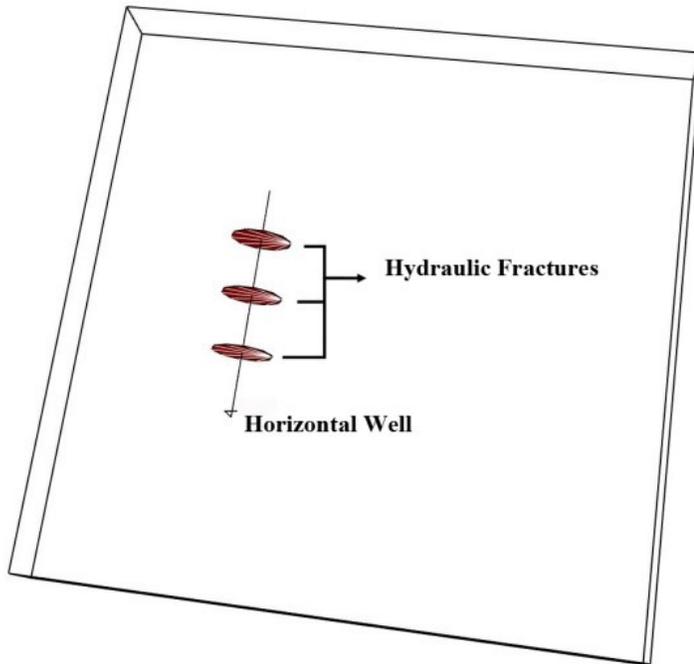
Figure 4-3 Well and hydraulic fracture locations.
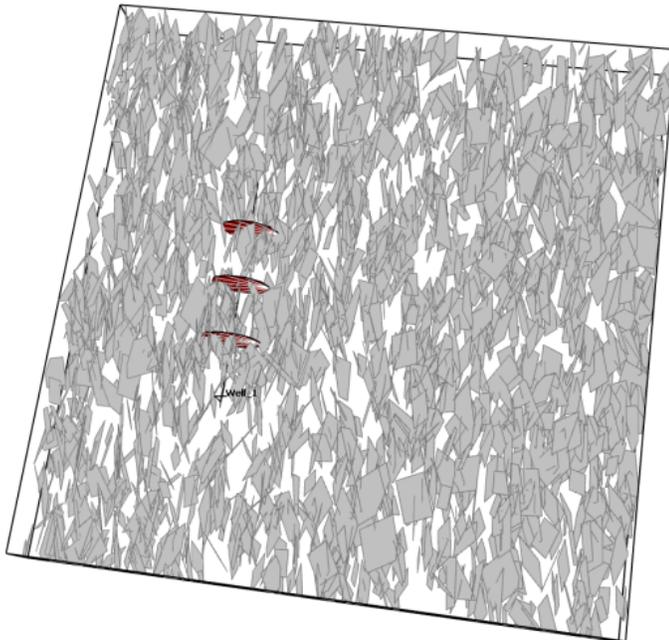
Figure 4-4 Fracture distribution of the true model.
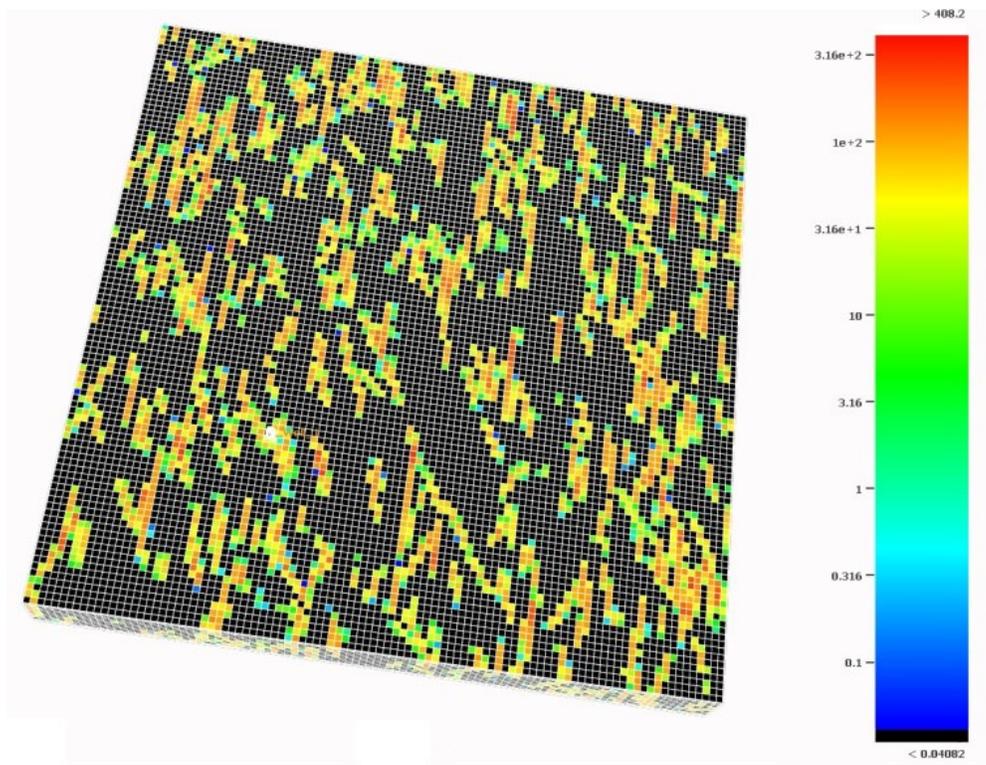
Figure 4-5 Upscaled fracture permeability in y direction (logarithmic scale), dual porosity reservoir simulation model.
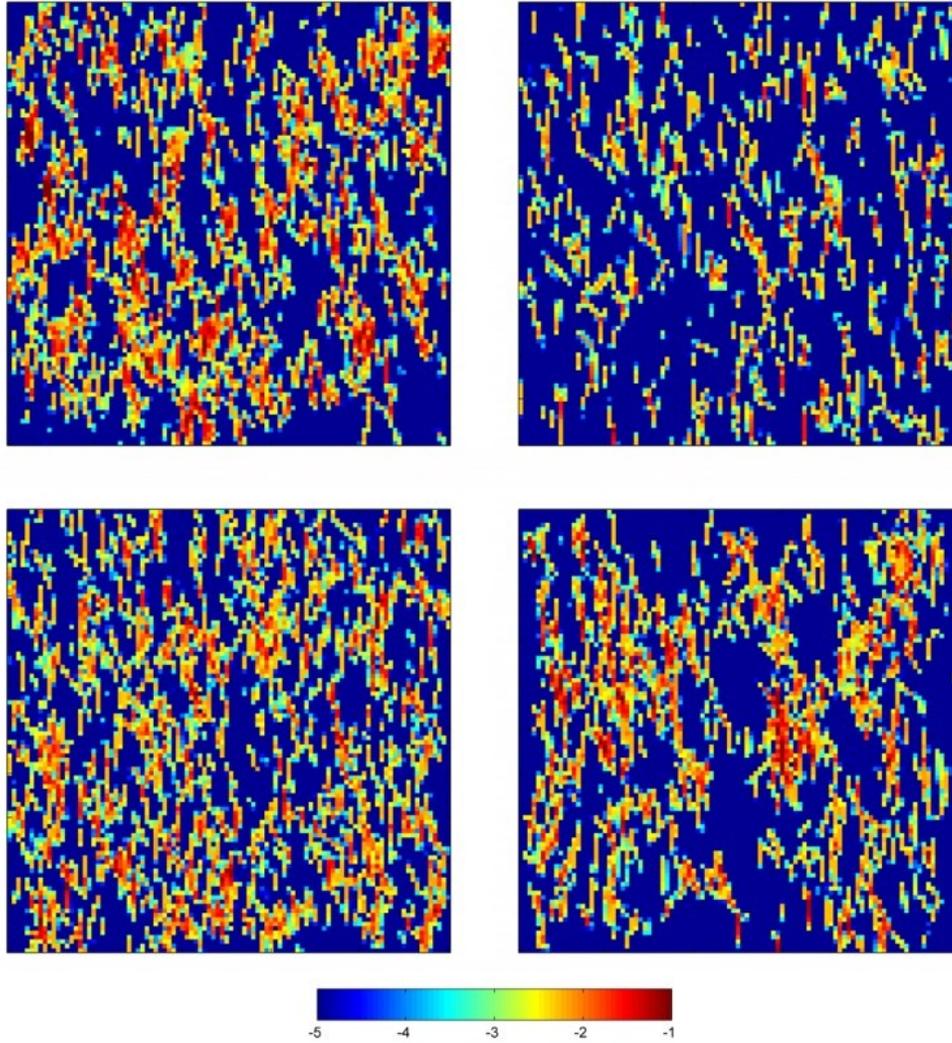
Figure 4-6 Fracture intensity map (P_{32L}) for some randomly selected initial ensemble members (logarithmic scale).
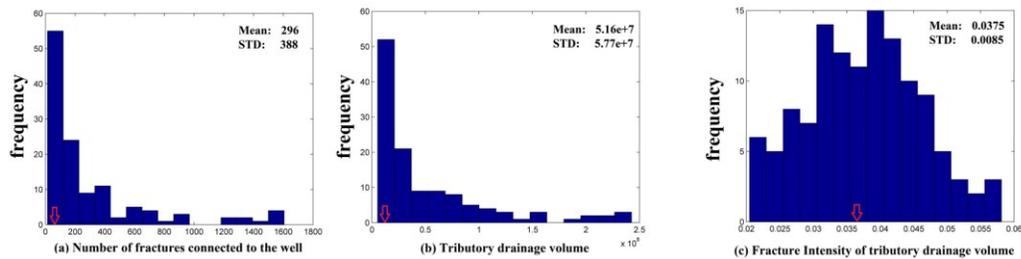
Figure 4-7 Histogram of the (a) total number of fractures connected to the well, (b) tributary drainage volume, and (c) fracture intensity of the tributary drainage volume for the initial ensemble.

145
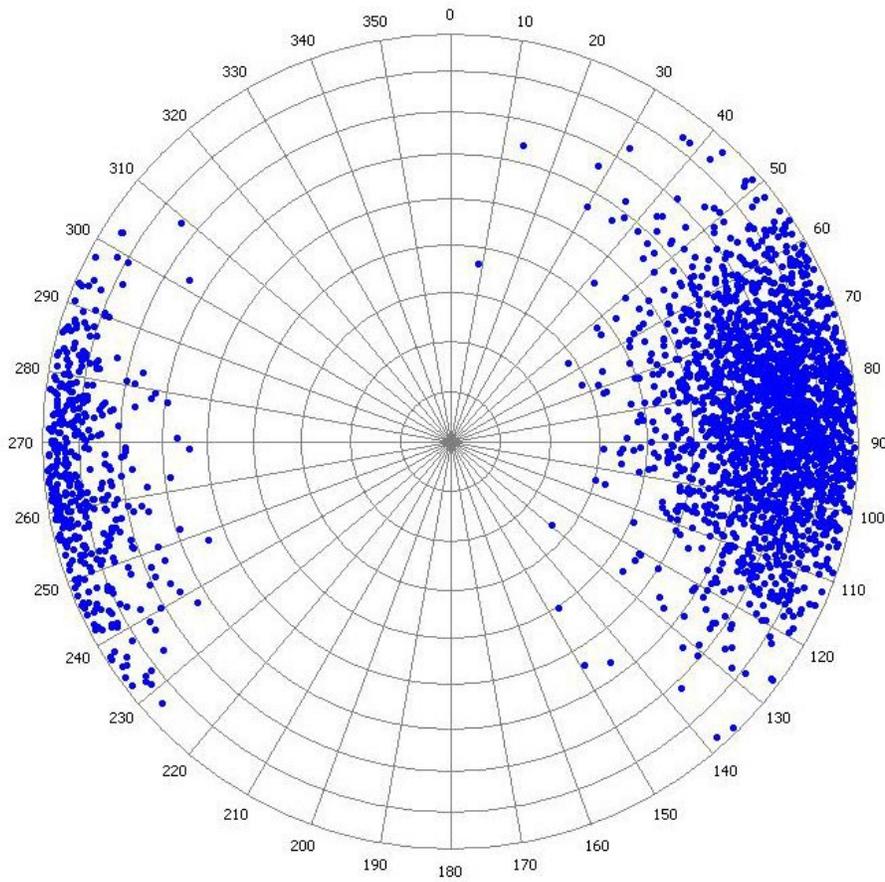
Figure 4-8 Stereonet illustration of fracture pole orientation for a randomly-selected initial ensemble member.
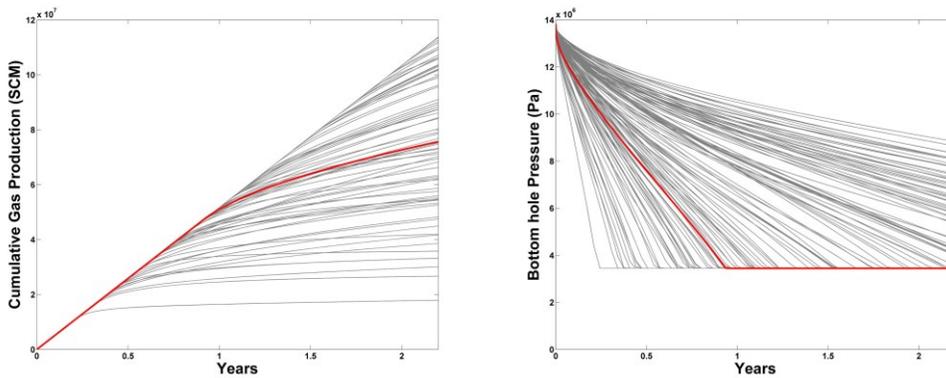
Figure 4-9 Cumulative gas production and bottom-hole pressure of the initial ensemble.

146
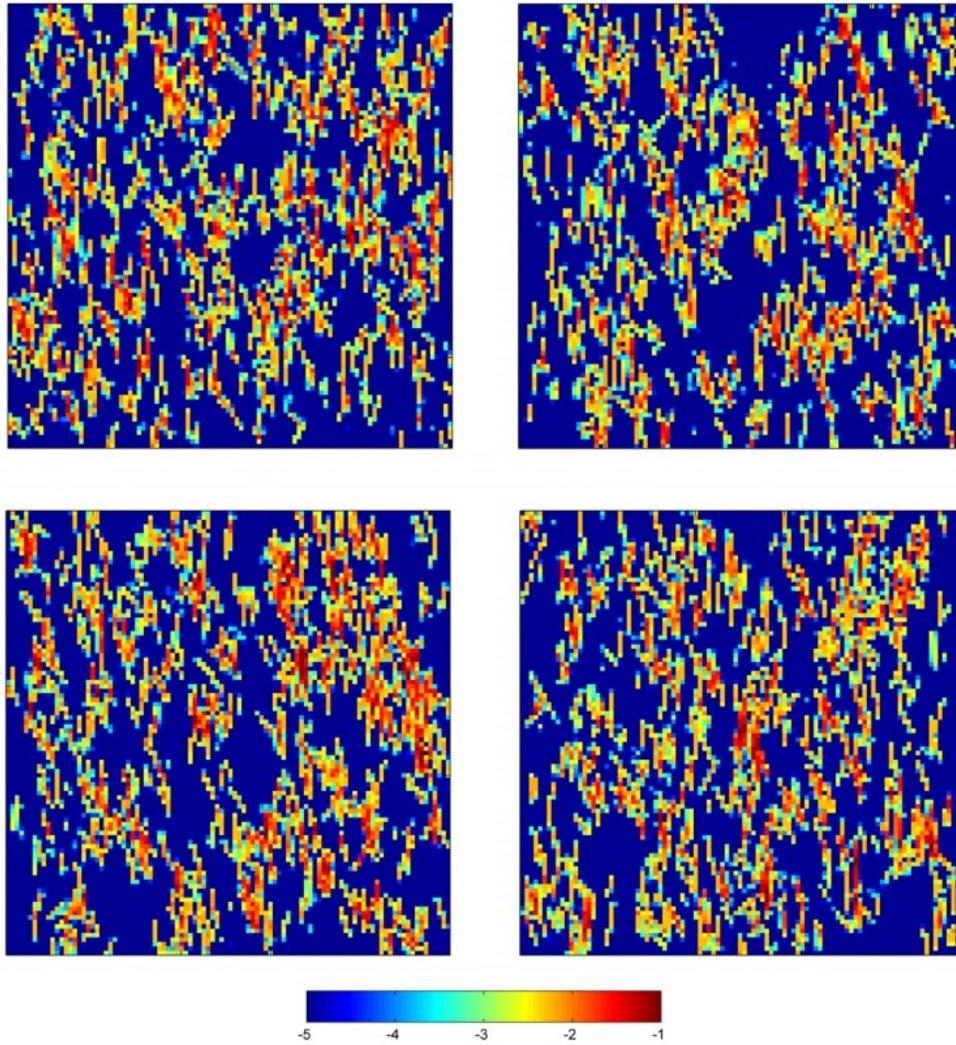
Figure 4-10 Fracture intensity map ($P_{32\,L}$) for some randomly selected updated models (logarithmic scale).
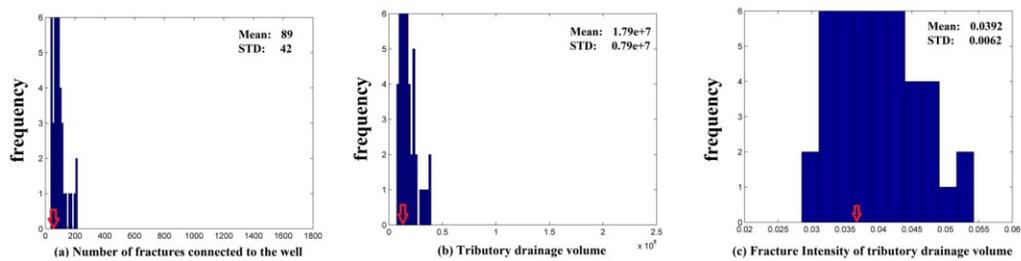
Figure 4-11 Histogram of the (a) total number of fractures connected to the well, (b) tributary drainage volume, and (c) fracture intensity of the tributary drainage volume ($P_{32\,G\text{-}TGV}$) for the updated realizations.
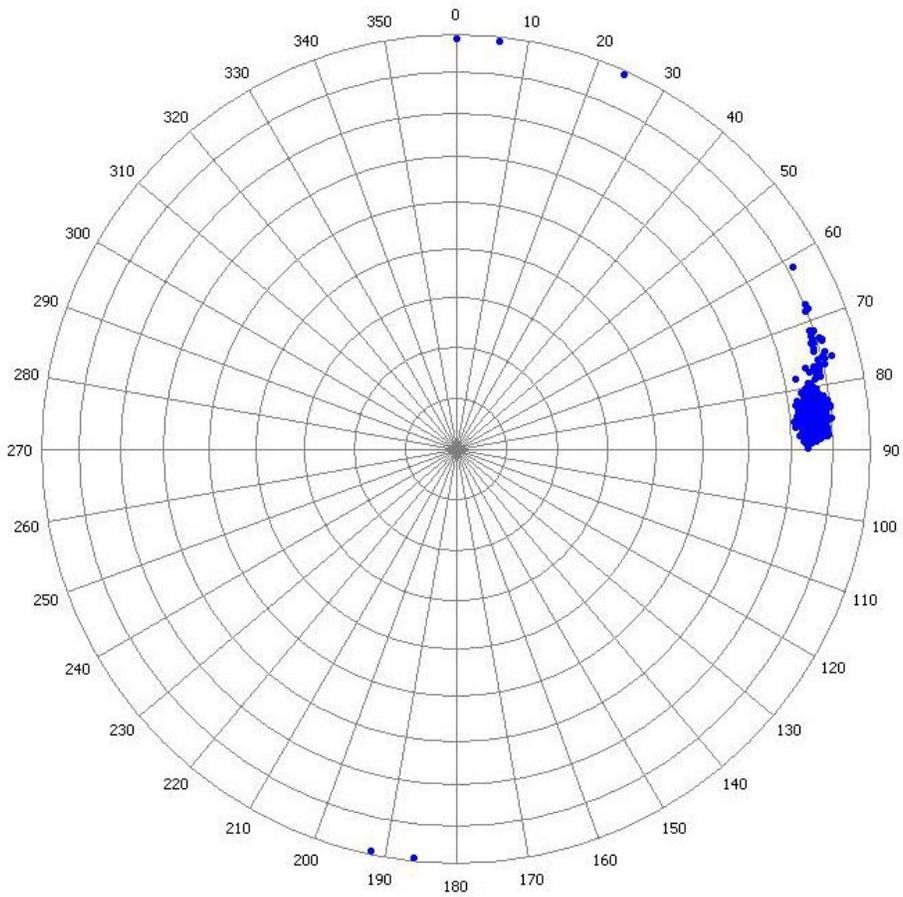
Figure 4-12 Stereonet illustration of fracture pole orientation for a randomly-selected updated ensemble member.
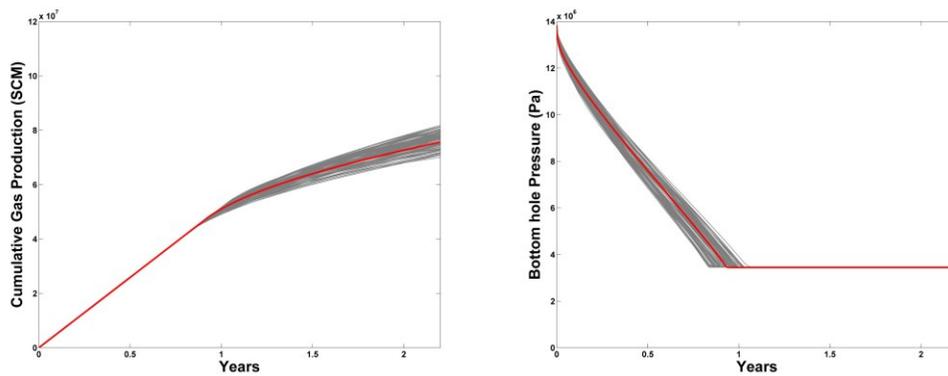
Figure 4-13 Cumulative gas production and bottom-hole pressure of the updated models.

Chapter 5: Integrated Characterization of Hydraulically Fractured Shale Gas Reservoirs with Microseismic Analysis, Rate Transient Analysis and Production History Matching[6]

5.1 Introduction

Shale gas reservoirs with low permeability rock matrix often experience uneconomical production rates. Successful (commercial) production requires maximum reservoir contact that is acquired through long horizontal well drilling and multi-stage hydraulic fracturing. Characterization of the fracture system, which facilitates fluid flow in the reservoir, is a key task in building a mathematical model for recovery predictions.

In conventional reservoirs, well test analysis is generally applied to characterize reservoir and fracture properties using various measurement data. Specific tests such as drawdown followed by a buildup period are conducted under controlled conditions to record downhole and surface flow measurements. In unconventional tight formations, the low reservoir permeability slows down the reservoir responses, and a relatively long testing time is required to measure the reservoir pressure transients, which is often not practical. Alternatively, analysis of dynamic production data, which are recorded over the life time of the well,

---

though might be less accurate compared to well test measurements, could provide valuable information for reservoir parameter estimations.

Production data analysis (PDA) has been widely adopted for quantification of reservoir and hydraulic fracture properties in tight gas and shale gas reservoirs. Many authors such as Clarkson (2013a, 2013b) have presented thorough reviews of the latest techniques for PDA and discussed workflows for production forecasting and reserve estimation. Analytical models, which are essentially simplified solutions to the detailed governing equations, have been employed extensively in pressure transient (PTA) and rate transient analysis (RTA) to estimate various system parameters such as hydraulic fracture half-length and fracture-matrix contact area (Bello, 2009). In the areas of PTA, Al Kobaisi et al. (2006) investigated the effect of fracture parameters using a hybrid model for a single fracture intersecting a horizontal well. Cheng et al. (2009) have used pressure data to estimate reservoir and hydraulic fracture parameters during the elliptical flow periods. They have estimated formation permeability, skin factor, and fracture half-length. Medeiros et al. (2008) have used a semi-analytical model to study the effects of natural and induced fractures as a dual porosity region together with the hydraulic fractures in shale gas reservoirs. In recent years, RTA has been widely adopted to characterize reservoir and hydraulic fracture parameters. Luo et al. (2014) have combined RTA with Monte Carlo analysis to assess uncertainties in matrix and hydraulic fracture properties. They demonstrated that uncertainties in matrix and fracture properties lead to uncertainty in production forecasts and estimated ultimate recovery.

150

Despite of the complex connectivity and configuration of a given hydraulic fracture system, analytical models used in most PDA methods would idealize the reservoir rock matrix and hydraulic fracture systems as simple domains by conceptualizing matrix as rectangular system, slabs, or other simple shapes with constant rock properties, separated by uniformly-distributed orthogonal fracture planes. This idealization is necessary to generate analytical or semi-analytical models representing fluid flow in the fractures and matrix-fracture fluid exchange. However, PDA or history matching is posed as an inverse problem with non-unique solutions. Analysis with analytical models would typically yield a homogeneous deterministic estimate representative of an average parameter value, but it fails to capture the effects of heterogeneity descriptive of the spatial variability in reservoir and fracture properties. In addition, analytical models often invoke assumptions including sequential depletion and a fully-connected fracture system, rendering detailed multiphase flow analysis in a multiple-porosity medium difficult. On the other hand, numerical simulations can be used to compute transient flow response in tight/shale gas media in which matrix blocks deplete into the fracture networks simultaneously within an arbitrary drainage volume (Alkouh et al., 2012).

RTA analysis is typically limited to some information about hydraulic fracture, mainly drainage volume, but fracture or SRV geometry cannot be obtained by these types of analysis. On the other hand, micro seismic monitory and analysis (MSMA) can provide information about limits of SRV as well as fracture geometry. MSMA is a rapidly evolving field that provides valuable information

151

for hydraulic fracture characterization and field development. It is used to determine the spatial extent of fractured rock resulting from stimulation and well treatment by mapping the locations of induced microseismic events. Fisher et al. (2005) have used microseismic analysis together with surface and downhole-tilt fracture mapping to characterize the induced fracture network. They have presented correlations between production responses and different fracture parameters. Typically, information from MSMA is used to constrain RTA and to select the appropriate model for analysis. There have been few studies in recent past comparing hydraulic fracture half-length estimates from different techniques such as RTA, PTA, microseismic and hydraulic fracture (fracture treatment pressure) modelling. Barree et al. (2005) and Clarkson (2011) provided such comparison for tight and shale gas reservoirs. It is important to note that these two studies provide valuable insight into estimates of fracture half-lengths by various methods and discrepancies between them. However to the best of the authors' knowledge, comparison of fracture properties estimated from RTA and MSMA with those obtained from detailed DFN models coupled with history matching has not been done. Besides, the history matched DFN models could potentially provide more information on effective fracture network than available through RTA, PTA, and MSMA.

In this chapter, production data, geologic information, and MSMA are integrated in an ensemble-based history matching technique to assimilate various DFN model parameters of the hydraulic fracture model. The result would be various DFN models and their equivalent dual continuum reservoir simulation models,

152

which honor both static geological data and dynamic information. In this field case application, the uncertain models parameters are hydraulic fracture radius and transmissivity as well as size, intensity, and transmissivity of the secondary induced fractures.

Numerous ensemble-based history matching techniques can be applied for practical production data integration (Emerick and Reynolds 2013). Ensemble Kalman filter (EnKF) is a Monte-Carlo based technique for data assimilation and has been introduced to the petroleum engineering industry by Lorentzen et al. (2001) as a promising approach for solving high-dimensional history matching problems (Aanonsen et al. 2009). An ensemble of initial models consistent with prior knowledge of the initial state and its probability distribution) are advanced to the time of the next observation using a forward model (e.g., numerical reservoir simulation). Prior joint probability distribution (covariance and mean of model states) are approximated directly from the ensemble of states and used to compute the Kalman gain, which is applied to update each ensemble member. The updated ensemble represents an empirical estimate of the posterior joint probability distribution of the model states (Liu and Oliver 2005). The primary assumptions are linear system dynamics and that the model parameters follow a multivariate Gaussian probability distribution function (pdf). Applying the updating schemes iteratively during each assimilation step has been proposed to address the issues related to non-linearities (Chen and Oliver 2012; Li and Reynolds 2009). In the case of DFN characterization, the highly nonlinear relationship between the DFN model parameters and the simulation model

production data implies that assimilating state variables (e.g. pressure data) is not practical and each proposed state requires a run of forward model after every update step to ensure consistency among state and model variables.   The Gaussian statistics assumption also presents an important challenge for applying EnKF to estimate fracture property distributions, which are commonly multimodal due to the underlying non-Gaussian geological features. A common strategy is to transform non-Gaussian model parameters from the original space to other space(s) in which they follow a nearly Gaussian distribution, such that EnKF can be applied in the transformed space. Various parameterization schemes such as truncated pluri-Gaussian, level set method, Gaussian mixture models, and discrete cosine transform have been proposed in the literature (Evensen 2007; Aanonsen et al. 2009). In this work, instead of characterizing individual fractures, five unknown model parameters (transmissivity and equivalent radius of the primary fracture, transmissibility, length, and global intensity of the secondary (induced) fractures), which are assumed to be constant and spatially invariant, are updated. These updated parameters are subsequently considered as conditioning data to populate the spatially-distributed secondary fractures based on a nearest neighbor model. Finally, each updated DFN model is upscaled to an equivalent dual porosity model for forward simulation modeling.

This chapter proposes an innovative methodology that incorporates numerical simulations with history matching workflow for parameter estimation of hydraulic fractures. The parameter optimization algorithm involves a probability

154

sampling step combined with ensemble Kalman filter (EnKF). The sampling technique, which is analogous to the use of importance weights in particle filtering, enables the algorithm to remove low weight models with a high probability. As the updated DFN models must be upscaled to equivalent dual porosity models prior to forward simulation modeling, additional nonlinearities between model parameters and system responses are introduced; retaining ensemble members with low weights leads to significant computational inefficiency and convergence difficulties.

The organization of this chapter is as follows: we begin our discussion with the formulation of the proposed EnKF and probability sampling procedures for dynamic data integration in DFN model parameter estimation. Next, we present the results obtained from the application of the proposed methodologies for a hydraulically-fractured shale gas model in Horn River Basin and the results are compared with production data analysis. At the end, the conclusion and remarks drawn from the case study are discussed.

5.2 Methodology

In this section, various components of the procedure are discussed in detail. The key techniques involved in the history matching procedure and optimization of discrete fracture network model parameters in hydraulic fractured wells are highlighted.

5.2.1 Generating discrete fracture network models

In this chapter, a hydraulic fracture is conceptualized as a fracture system consisting of a primary fracture, which is an elongated penny-shaped crack, together with secondary fractures, which are complex induced fractures connected to the primary fracture. This definition is consistent with the fracture system described by Fisher et al. (2005) as complex or very-complex hydraulic fracture system.

Generating discrete fracture network models involves defining basic features of the fractures including intensity, orientation, location, size, elongation, and other fracture properties such as aperture and transmissivity, which are defined individually for each fracture set in the model. Initially, features associated with the primary fracture are specified. In this work, fracture location, orientation, and aperture are considered to be constant and known, while fracture transmissivity (T^{pf}) and equivalent radius (R^{pf}) are unknowns. Subsequently, secondary fractures are defined in relation to the primary fracture. Locations of the secondary fractures are defined by means of nearest neighbor model, in which the local fracture intensity (P_{32L}^{sf}) decreases exponentially with distance from the primary fracture. The intensity (P_{32L}^{sf}) at any point in space is defined by

$$P_{32}(x) = ce^{-b\delta},$$

5-1

where c and b are constants and δ is the distance between location x and the primary fracture plane (Dershowitz 1993). The orientation parameters of the induced fractures are defined using Fisher distribution (Fisher 1953). The global

156

fracture intensity (P_{32G}^{sf}) of the induced fractures, a scalar parameter equivalent to the total number of fractures in the model (fracture count), is considered as an uncertain parameter in the case study. Fracture length (L^{sf}) and transmissivity (T^{sf}) are the other unknown parameters of the secondary induced fractures.

Initial ensemble of discrete fracture network models is generated by defining primary and secondary fractures, conditioned to the known parameters and considering specific distributions for uncertain parameters. The uncertain parameters are updated during history matching and an improved ensemble of DFN models are constructed from the assimilated/updated parameter values using a fracture generation algorithm as implemented in FRACMAN® (2011).

5.2.2 Upscaling discrete fracture network models

Once a discrete fracture network model is generated, it is translated into dual porosity reservoir simulation model, which is an equivalent representation of the fracture network and the matrix systems defined over a structured mesh. In this implementation, Oda's (1985) analytical approach is used. Oda's technique is a static approach without any flow simulation computations. This method is fast and efficient, and it is well suited for well-connected and high-density fractures; this assumption is practical for upscaling primary and secondary induced hydraulic fractures that are by definition, highly connected to each other.

5.2.3 Forecast model

In a fractured reservoir, fluids exist in two connected system; the matrix and the fractures. Typically fractures provide high conductivity conduits and the matrix

provides the fluid storage. A fractured reservoir with relatively low matrix permeability, but high porosity, may be characterized by a dual porosity/single permeability model.

The forecast model is separately applied to each ensemble member. The multiphase fluid flow in petroleum reservoirs is described by material balance, momentum balance, phase behavior descriptions and numerous auxiliary equations. This system of non-linear differential equations is solved numerically using methods such as finite difference

$$\begin{bmatrix} u_k^p \\ d_k^p \end{bmatrix} = G \begin{bmatrix} m_{k-1}^a \\ u_0^a \end{bmatrix},$$

5-2

where u denotes the estimated state variables (pressure and saturation) and d is the production data predictions, after running the dual porosity dual permeability reservoir simulator (forward model, G) from the initial conditions at time 0 to time step k. n_e is the total number of ensemble members and the superscripts, p and a, specify the predicted and analysed states, respectively..

5.2.4 Probability sampling of the proposals

Considering the history matching problem in the Bayesian framework, it is desired to find the maximum a posteriori estimate (MAP) of discrete fracture network model parameters m. Observed data d_{obs} is used as the conditioning information to estimate the posterior probability density function (pdf) that is given by

$$p(m \mid d_{obs}) = L(d_{obs} \mid m) = a \exp(-O(m))$$

5-3

where $O(m)$ denotes the objective function

$$O(m) = \frac{1}{2}[G(m) - d_{obs}]^T C_D^{-1}[G(m) - d_{obs}] + \frac{1}{2}[m - m_{prior}]^T C_M^{-1}[m - m_{prior}]$$

5-4

Here, the objective is to sample realizations based on their likelihood (Equation 5-3). The likelihood function is used to assign a weight to each realization; all proposals have a probability to be accepted and promoted to the next step. This techniques, enables the algorithm to explore the solution space and to avoid local minima. Ma et al. (2008) have proposed a two-stage Markov Chain Monte Carlo (MCMC) method for quantification of permeability uncertainty in history matching reservoir models. They start by computing the observed data mismatch for a proposed change in model parameter based on a linearized approximation to flow simulation. If the proposed changes satisfy the acceptance criterion and the estimated objective function match is better than that calculated for the current state in the Markov Chain, the proposals are subjected to full flow simulation. Using this two-stage algorithm to pre-screen the proposals, the computational costs are reduced. Emerick and Reynolds (2012) suggested an Ensemble Kalman filter Markov Chain Monte Carlo procedure where MCMC is implemented to generate the desired number of samples for EnKF. They proposed a sampling based on the objective function. They have reweighted the Markov Chain samples based on the likelihood and implemented the concept of importance sampling to remove low probability

159

samples. They proposed a weighting function that is proportional to $\exp\left(-O(m)\right)$. Nejadi et al. (2014) have also successfully adopted an analogous probability weighting scheme in the EnKF framework to maintain the reference statistics of the model parameters and to improve ensemble variability.

In this study, a weighting scheme similar to that of Emerick and Reynolds (2012) is proposed. The likelihood of the model parameters m is defined based on the production mismatch.

$$p\left(m \mid d_{obs}\right) = a \exp\left(-\frac{1}{2}[G(m) - d_{obs}]^T C_D^{-1}[G(m) - d_{obs}]\right).$$
5-5

The implementation in Equation 5-5 is similar to the concept of importance weights in particle filter technique, where the importance weight w_j of ensemble member j is computed as

$$w_j = \frac{\pi(m_j)}{\displaystyle\sum_{k=1}^{n_e} \pi(m_k)},$$
5-6

where $\pi(m_j)$ is defined as

$$\pi(m_j) = \exp\left(-O(m_j)\right) = \exp\left(-[G(m_j) - d_{obs}]^T C_D^{-1}[G(m_j) - d_{obs}]\right).$$
5-7

The objective function in Equation 5-7 should be defined based on the nature of the history matching problem. If prior information regarding the model parameters is available, the objective function can be modified to capture the mismatch in model parameter statistics. The sampled members are used as an input to EnKF algorithm. With this sampling procedure, all proposed realizations

160

have a certain selection probability, avoiding getting trapped in local optima. In addition, the proposals with large mismatch (high value of $O(m)$) would have low importance weights and are more likely to be removed. This selection approach is analogous to the use of importance weights in particle filtering techniques, where ensemble members (particles) are weighted on the basis of likelihood function. Ensemble members are promoted based on their likelihood.

5.2.5 Analysis step – Ensemble Kalman filter

The analysis step is carried out using the observation data (d_k) for the selected samples

$$y_k^{a,j} = y_k^{p,j} + KG_k \left(d_{ik} - Hy_k^{p,j} \right),$$

5-8

where $y_k^{a,j}$ and $y_k^{p,j}$ represent the analysed state vector and the posterior state vector of ensemble member j, respectively; K_k is the Kalman gain at k^{th} step, and H is the observation operator, which represents the relationship between the state vector and the observation vector

$$H = \begin{bmatrix} 0 | I \end{bmatrix},$$

5-9

d_k^j is the summation of observed production data at the kth step ($d_{obs,k}$) and ε_k^j, a vector of measurement errors such that $E\left[\varepsilon \varepsilon^T \right] = C_D$

$$d_{ik} = d_{obs,k} + \varepsilon_k.$$

5-10

Kalman gain K_k is defined as

$$K_k = C_{y_k^{p,j}} H^T \left(H C_{y_k^{p,j}} H^T + C_{d_k} \right)^{-1},$$

5-11

where $C_{y_k^{p,j}}$ is the state cross covariance matrix, and C_{d_k} is the error covariance

matrix.

The cross covariance matrix is approximated as

$$C_{y_k^{p,j}} \approx \frac{1}{N_e - 1} \left\{ \left(y_k^p - \overline{y_k^p} \right) \left(y_k^p - \overline{y_k^p} \right)^T \right\},$$

5-12

where T indicates matrix transpose, and $\overline{y_k^p}$ is the average of all posterior state

vectors.

In this study, the state vector of the ensemble Kalman filter formulation (y_k^j)

consists of the DFN model parameters (m_k^j) and the production data predictions

(d_k^j), which are the simulation outputs of the dynamic model

$$y_k^j = \begin{Bmatrix} m_k^j \\ d_k^j \end{Bmatrix}.$$

5-13

The following parameters are included in the state vector as the model

parameters and updated using EnKF algorithm:

- Transmissivity of the primary (elongated penny-shaped) fracture (T^{pf})

- Equivalent radius of the primary (elongated penny-shaped) fracture (R^{pf})

- Global fracture intensity of the secondary (induced) fractures (P_{32G}^{sf})

- Transmissivity of the secondary (induced) fractures (T^{sf})

- Length of the secondary (induced) fractures (L^{sf})

The model parameters (m_k^j) is a 5×1 vector for different realizations

$$m_k^j = [T^{pf}, R^{pf}, P^{sf}_{32G}, T^{sf}, L^{sf}]^{T\,j}_k \,.$$

5-14

For each updated member of the selected ensemble, new models are generated such that the actual ensemble size remains unchanged. In addition, this re-sampling step can help to avoid ensemble collapse and maintain diversity among the re-sampled ensemble members.

5.3 Field Case Study

The objective of this case study is to illustrate the history matching workflows described in the Methodology for characterization of a hydraulically-fractured shale gas well in the Horn River basin. It further demonstrates the applicability of the proposed framework for uncertainty quantification of hydraulic fracture parameters for shale gas reservoirs.

The application case is taken from the Horn River Resource Play, located approximately 100 km northeast of Fort Nelson, BC, Canada. The formation under development is the Devonian age Horn River Group comprised of the Muskwa, Otter Park, and Evie shales. Four-stage hydraulic fracturing operation is used for well treatment and improving its productivity.

5.3.1 Model description

The dimensions of the model are approximately $750 \times 900 \times 175$ m^3. It consists of a single multi-fractured horizontal well. Microseismic data is acquired for hydraulic fracturing treatment of the well. Figure 5-2 shows the recorded

microseismic activity locations. The approximate trend (the horizontal angle in the x-y plane away from the North) and plunge (the vertical angle in space away from horizontal x-y plane) of the primary and induced fractures of each stage are inferred from the microseismic responses. Figure 5-3 shows the pole orientation (orientation of the vector normal to the fracture plane) for the primary and induced fracture planes. MSMA has determined that the fracture growth occurs predominantly along the northeast direction of the treatment zone. The approximate azimuth angle of the fractures is 50 degrees measured from North (equivalent pole trend of 140 degrees) and 600 out of 900 events are contributing to the effective fracture volume.

Roughly 12 months of production history is available for the well. Four months of the production history is used for parameter optimization and history matching; the rest of the history is used to assess the production forecasting capability of the updated models. CMG Gem simulator (2013) is used as the reservoir simulator (forward model). Table 5-1 summarizes the parameters of the discrete fracture network model.

5.3.2 Workflow

Implementation details of the described workflow in this field case study are presented here. Initial probability distributions of the uncertain model parameters (Equation 5-14) are modeled as Gaussian random fields. In particular, spatial distribution of the micro seismic point data is applied to infer the approximate statistical parameters of fracture size. The mean of R^{pf} and L^{sf} are estimated as

97 and 50 meters, respectively. The probability distributions of T^{pf}, T^{sf} and P^{sf}_{32G} are constructed to encompass the entire possible range of uncertainties, such that true value would fall in between the distribution minimum and maximum. Additional geologic information could be helpful to refine this range. Table 5-2 and Figure 5-4 summarize the distribution statistics of the five unknown model parameters. Fifty random vectors of the model parameters ($n_s = 50$) are sampled from these initial probability distributions. Next, three specific ensemble members ($n_d = 3$) are generated corresponding to each random vector of the model parameters in a Monte Carlo simulation. Figure 5-5 illustrates an example of initial discrete fracture network model that is generated from the average values of the initial probability distributions.

The initial ensemble of DFN realizations are subsequently upscaled into their equivalent dual continuum geocellular models using Oda (1985) analytical upscaling technique. The geocellular model consists of a $15 \times 18 \times 14$ grid. Cartesian local grid refinement is used to enhance the grid definition in the areas near the horizontal well and to improve numerical accuracy. A two-step approach is performed to upscale the near-wellbore regions over a locally-refined grid and the rest of the reservoir over the global grid as defined previously. Parameters of the upscaled dual porosity simulation model are summarized in Table 5-1. Petrophysical properties (porosity and permeability) of the matrix system are assumed to be constant for each layer. Figure 5-1 shows the dynamic model structure and porosity values of individual layers for the

matrix system. The upscaled dynamic models ($n_s \times n_d = 150\,\mathrm{models}$) are subjected to the forward model for the first production period ($t_0 = 0$ to $t_1 = 15\,days$). Figure 5-6 shows the gas production rate of the initial ensemble compared to the actual gas production rate measurements.

In order to assess the performance of individual realizations, an objective function based on production mismatch is implemented to assign a weight to each ensemble member, as explained in Equation 5-6. The weighting factors are used in a probability sampling approach to randomly select n_s members, which are considered as the prior in EnKF. Taking into account the data likelihood, posterior members are obtained from EnKF update. For each of the n_s updated members, n_d new models are generated. As discussed in the Methodology section, this re-sampling step helps to ensure constant ensemble size, avoid ensemble collapse, and maintain diversity among the re-sampled ensemble members. Unlike traditional EnKF, where Bayesian update is combined with advancing the model in time, as new data becomes available, the posteriors are used to generate new DFN models, which are then subjected to flow simulation from t_0 to the next assimilation time step t_2. This type of iterative EnKF scheme is often implemented to ensure consistency between model parameters m and data d when G(m) is a highly non-linear function.

The prescribed workflow is repeated until the last assimilation time step to incorporate the entire production history. In this study, history matching is performed for 5 steps using 4 months of the production history. In each step the

objective function is calculated for the time period $t_0 = 0$ to t_k, where t_k is equal to 15 days for $k = 1$ and is equal to 1, 2, 3, and 4 months for $2 \leq k \leq 5$. Figure 5-7 shows a randomly-selected discrete fracture network model from the updated ensemble, and Figure 5-8 shows the gas production profiles of the updated models compared to the true production values. Also presented in Table 5-2 are the distribution statistics of the updated DFN model parameters.

5.3.3 Rate transient analysis

RTA is also performed to estimate numerous reservoir and fracture parameters including stimulated reservoir volume, matrix or fracture permeability, fracture half-length and contacted matrix-fracture area. The analysis models implemented in Fekete Harmony ^{TM} (2014) and F.A.S.T. RTA ^{TM} software are summarized below:

- Flowing material balance (Moghadam et al. 2011)

- Analytical models (Brown et al. 2011)

    - Horizontal multifrac composite model

    - Horizontal multifrac SRV (Uniform) model

- Type curve

    - Agarwal-Gardner (Agarwal et al. 1999)

    - Blasingame (Anderson et al. 2006, Amini et al. 2007)

    - Normalized pressure integral (Blasingame et al. 1989)

    - Transient (Finite conductivity fracture)

    - Wattenbarger (Wattenbarger et al. 1998)

- Numerical

  - Horizontal multifrac model

- Unconventional reservoirs

Model assumptions and type curve analysis for selected techniques are summarized in Figure 5-9. Table 5-3 compares the results of different analysis methods with those obtained from the history matching procedure presented in this chapter. In some cases, the production data does not follow a particular type curve but scatter over a number of curves; the plots in Table 5-3 are chosen to represent the best attainable match and the corresponding model parameters derived from the analysis.

5.4 Results and discussion

In the field case example, we examine the ability of the proposed algorithm to characterize a number of hydraulic fracture parameters. The initial realizations are created based on the available geologic and microseismic data, capturing the uncertainty in the prior model. Figure 5-6 illustrates the production profiles of the initial ensemble and the proximity of an individual member to the true dynamic response. Although the true production history is captured among the initial ensemble, it is clear that the uncertainty exhibited by the initial ensemble is too large. The high production mismatch reflects the uncertainty in the initial ensemble. The objective function decreases as the history matching progresses. It is interesting to note that although only a small portion of the production history has been integrated in the model assimilation workflow (4 months), capability of the updated ensemble to predict the future well performance

168

(remaining 8 months of available production data) is remarkable, as evident in Figure 5-8.

Static data such as well logs and microseismic interpretations are typically insufficient to infer all relevant parameters of the hydraulic fracture systems. For example, secondary induced fracture parameters such as intensity and transmissivity cannot be directly obtained from microseismic analysis. However, it is demonstrated that utilization of dynamic flow data could extend the observability of the model parameter estimation by integrating complex fluid flow physics. The history matching procedure also assists the proper quantification of uncertainties in model parameters. Figure 5-4 compares statistics of the five unknown parameters for the initial and updated ensemble; the same statistics are also summarized in Table 5-2. For all model parameters, higher variances (wider ranges) are observed among the initial distributions, which are derived based on microseismic estimates or prior knowledge extracted from nearby fields. The history matching process has refined these distributions conditioning to both dynamic and static information. As a result, variance exhibited by the updated ensemble is diminished, reflecting a reduction in uncertainty in the updated models. Mean of the updated ensemble is also adjusted. For example, radius of the primary hydraulic fracture plane is significantly reduced (Mean: 19.76 and STD: 18.55) compared to the initial microseismic analysis (Mean: 97.55 and STD: 36.22). This is because the hydraulic fracture size interpreted from microseismic analysis represents created fracture, whereas the effective fracture network available to flow is much

smaller. It is believed that the production data usually give pessimistic but accurate measure of the hydraulic fracture parameters due to effective flow. These observations are in-line with those made by Baree et al. (2005) and Clarkson (2011) for tight gas reservoirs. Clarkson (2011) observed that well-test, frac-modelling and RTA-derived values were lower than miscroseismic-derived values due to created nonconductive fractures.

Various PDA models have also been employed to estimate the hydraulic fracture half-length, stimulated reservoir volume, and effective reservoir permeability. It is clear that different models often produce different results depending on the assumptions and ensuing simplified analysis equations. For example, a wide range of effective permeability estimates (e.g., near-wellbore areas surrounding the hydraulic fractures and outer or matrix regions) can be obtained using a number of different models. In the updated models obtained from the history matching procedure, locations, sizes, and transmissibilities of the primary and secondary induced fractures are considered, providing a more comprehensive characterization of the entire hydraulic fracture system. Finally, it is noted that the estimated primary fracture half-length (20m) is within the observable range of 6.71 to 98 m from PDA. Similar conclusions can be derived by comparing the stimulated reservoir volume (9e+6) to the wide range of 3 to 105e+6 m3 from PDA (Table 5-3).

The final updated models take into account effective conductivities and contact area between matrix and fracture systems at different scales (primary and secondary). Another advantage of the proposed approach is that it can be readily

170

adopted, if necessary, to estimate additional unknown parameters. For example, fracture aperture, which is generally unattainable from other means including PDA, can be considered as an unknown parameter. The algorithm presented in this study provides a suite of equally probable equivalent dual-media models and their equivalent DFN models that are useful and reliable for future production forecasting.

5.5 Conclusion

Estimation of hydraulic fracture parameters is often challenging because reservoir characterization based on both static and dynamic data is an inverse problem that is highly nonlinear and the solutions are not unique. In this chapter, a novel procedure is proposed for dynamic data assimilation in hydraulically-fractured reservoirs. The method entails combining probability weighted sampling with Ensemble Kalman filter. A weighting scheme based on dynamic data mismatch is proposed. Results from a field case study demonstrate the applicability of the described technique in characterizing a number of hydraulic fracture parameters for a shale gas model. Multiple discrete fracture network models and their equivalent dual media are updated. It is also illustrated how the proposed history matching procedure can be integrated with microseismic analysis to obtain realistic estimates of hydraulic fracture parameters along with the corresponding uncertainties.

In order to ensure consistency between model and data vectors, an iterative scheme is implemented where after every update step, the updated ensemble members are subjected to the forward model to compute the dynamic responses

from the beginning to the next assimilation step. Although additional computational efforts are incurred with these extra forward model executions, the improvement in estimation accuracy is significant enough to offset the extra costs. Therefore, the proposed procedure presents a promising potential for history matching and uncertainty assessment of hydraulic fracture model parameters.

Table 5-1 Parameters of the discrete fracture network model and the dynamic dual porosity simulation model.

Discrete fracture network parameters:		
Primary Fractures		
Aperture	1e-3 m	
Storativity	1 e-6 1/m	
Trend (Stage#1)	226 °	
Trend (Stage#2)	357 °	
Trend (Stage#3)	64 °	
Trend (Stage#4)	68 °	Obtained from Microseismic
Plunge (Stage#1)	19 °	
Plunge (Stage#2)	-68 °	
Plunge (Stage#3)	-28 °	
Plunge (Stage#4)	-55 °	
Secondary Fractures		
Aperture	1e-6 m	
Storativity	1 e-6 1/m	
Trend	Similar to primary fracture	
Plunge		
Fisher orientation distribution	10	
Dynamic model specifications:		
Model dimensions	$750 \times 900 \times 175$ m^3	
Number of grids	$15 \times 18 \times 14$	
Reservoir depth	1731 m	
Matrix porosity	5 to 7 %	
Matrix permeability	2.61 e-4 to 3.56 e-4 md	
Initial reservoir pressure	33,000 kPa	

Table 5-2 Mean and standard deviation (STD) of different discrete fracture network model parameters for the initial ensemble compared with the updated realizations.

Parameter	Initial Models	Updated Models
Transmissivity T^{pf}	Mean: 280.88 STD: 120.26	Mean: 339.56 STD: 44.41
Radius R^{pf}	Mean: 97.55 STD: 36.22	Mean: 19.76 STD: 18.55
Intensity P_{32G}^{sf}	Mean: 0.0069 STD: 0.0013	Mean: 0.0091 STD: 0.0005
Transmissivity T^{sf}	Mean: 31.16 STD: 18.70	Mean: 44.91 STD: 4.68
Length L^{sf}	Mean: 49.58 STD: 15.83	Mean: 44.42 STD: 8.65

Table 5-3 Fracture half-length, stimulated reservoir volume and formation permeability obtained from different production analysis techniques compared with our methodology.

	Analysis Name	Fracture Half-Length (m)	Stimulated Reservoir Volume (m3)	Permeability (md)
Production Data Analysis	FMB-Gas	-	3,492,841	-
	Gas-AM-Hz Multifrac Composite	6.71	105,000,093	0.00789 Inner Zone 0.00020 Outer Zone
	Gas-AM-Hz Multifrac Uniform	6.71	3,523,103	0.02694 Inner Zone 0.01811 Outer Zone
	Gas-AG-Fracture	11.54	7,325,771	0.00742
	Gas-Blas-Elliptical	42.89	8,364,088	0.00549
	Gas-Blas-Finite Cond.	11.72	7,550,301	0.01047
	Gas-Blas-Fracture	16.24	3,623,978	0.00514
	Gas-Blas-Horizontal	-	4,604,596	-
	Gas-NPI-Fracture	9.08	4,536,670	0.00802
	Gas-Transient-Finite Cond.	98.01	6,723,621	0.00145
	Gas-Wattenbarger	22.89	2,963,834	0.01519
	Numerical-Hz Multifrac	13.20	105,000,093	0.00351 Inner Zone 0.00020 Outer Zone
	UR-GST	6.71	3,523,103	0.01811 Matrix 1.30E-05 Effective
	UR-GVPH	16.31	8,561,218	0.00315 Matrix 1.30E-05 Effective
DFN History Matching		19.76 Main crack 65 Main+Induced	8,909,179	-
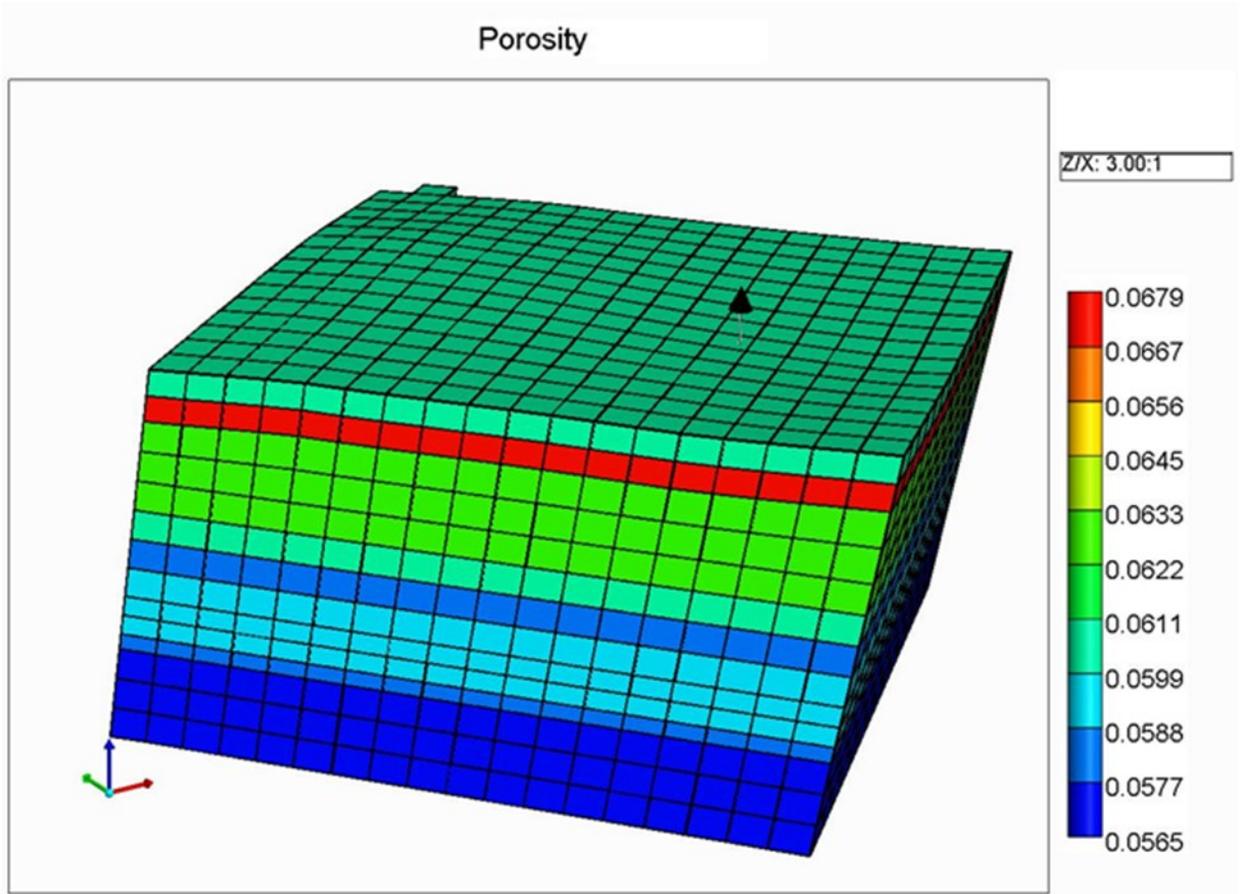
Figure 5-1 Grid structure and matrix porosity of the dynamic reservoir model.
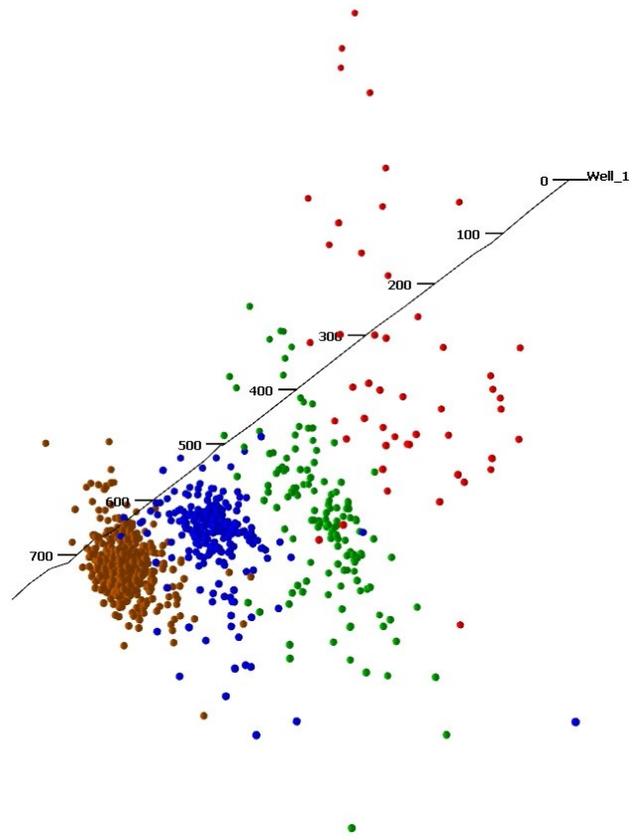
Figure 5-2 Microseismic event locations (brown: stage 1; blue: stage 2; green: stage 3 and red: stage 4)
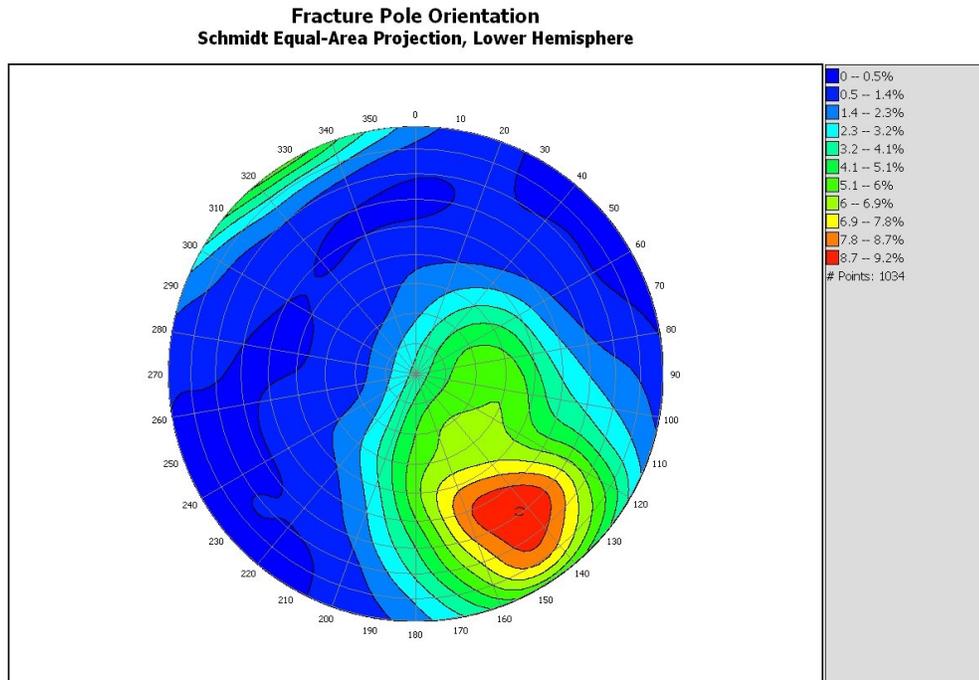
Figure 5-3 Contour plot showing the fracture pole orientation of primary and secondary (induced) fractures from microseismic analysis.
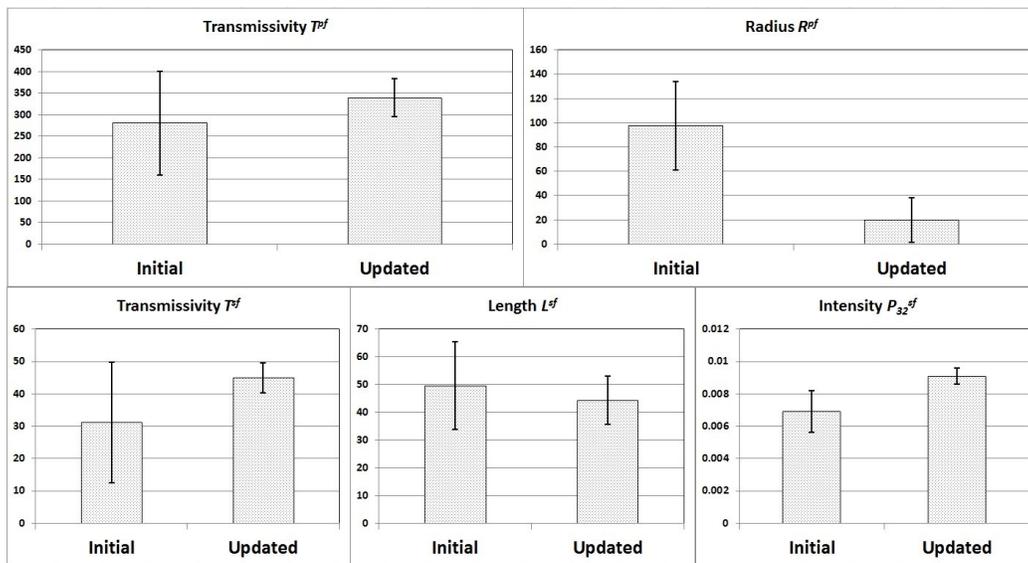
Figure 5-4 Comparison of primary and secondary fracture parameter statistics for the initial and updated ensembles. The error bar represents variation around the parameter mean.

176
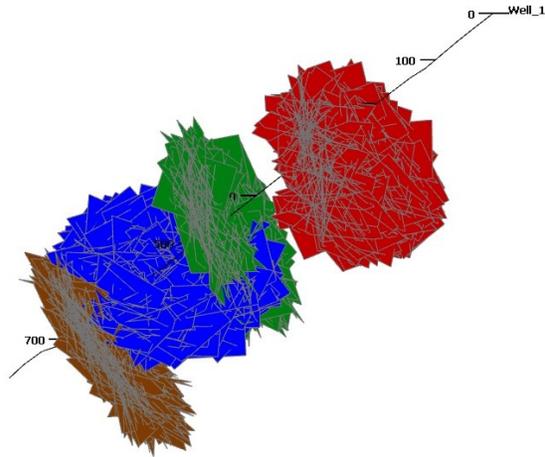
Figure 5-5 Illustration of an initial discrete fracture network model based on parameter distribution means.
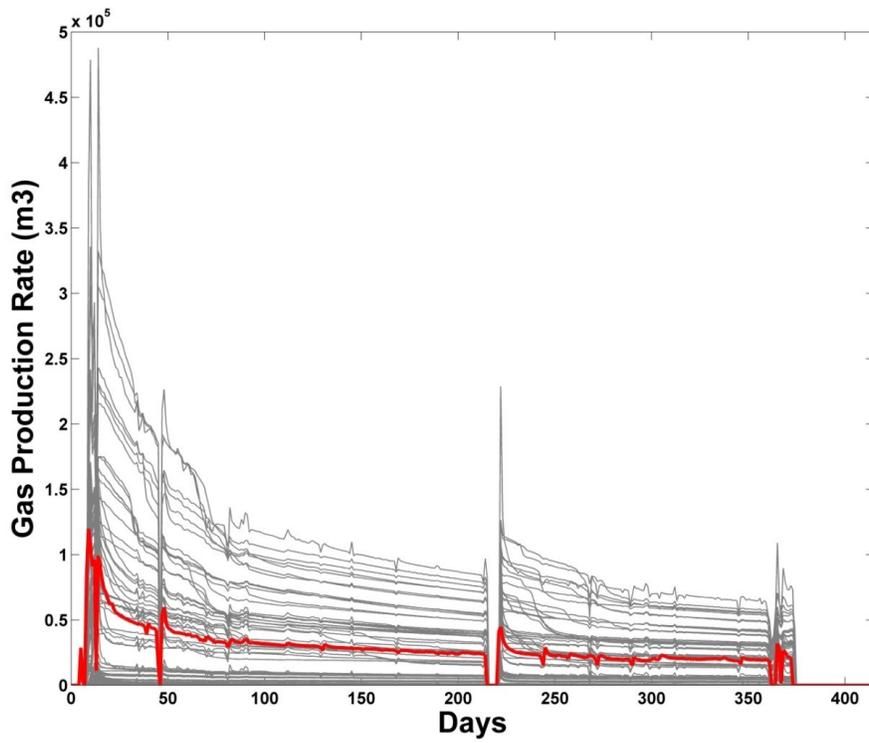
Figure 5-6 Gas production profiles of the initial ensemble compared to the observed production data.
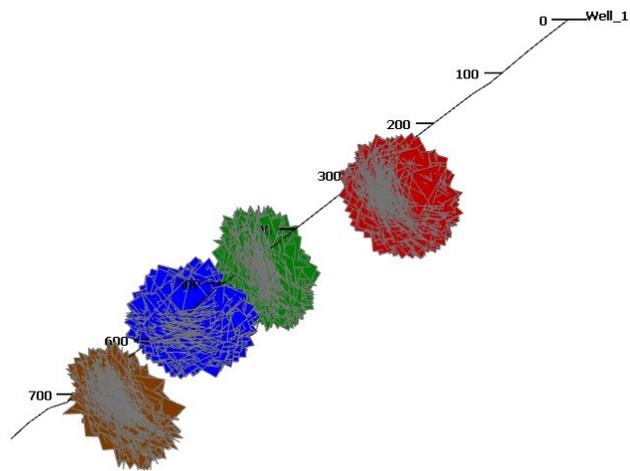
Figure 5-7 Illustration of a discrete fracture network model after EnKF updating.
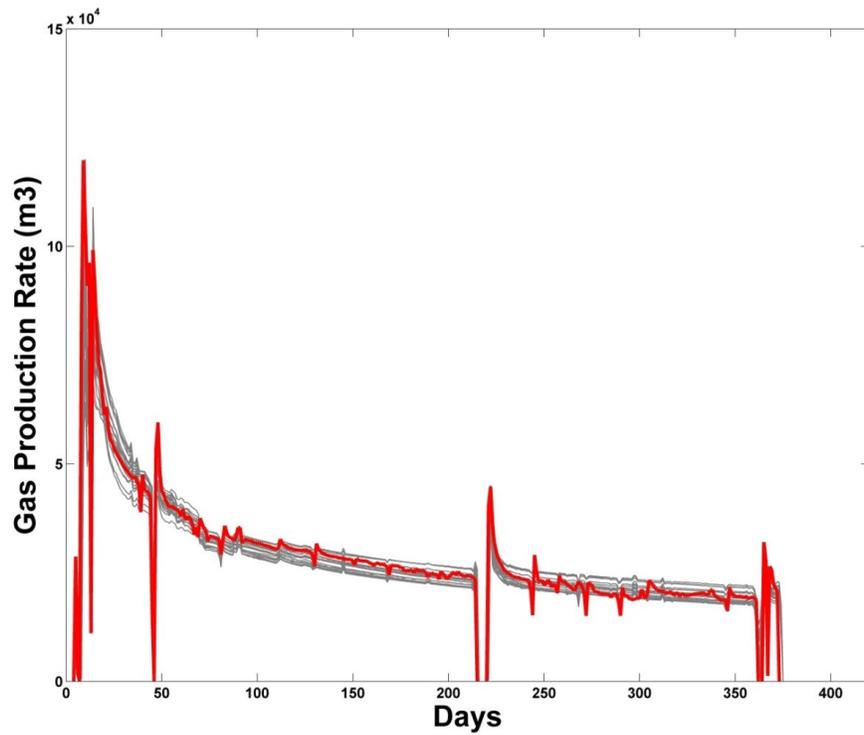
Figure 5-8 Gas production profiles of the updated models compared to the observed production data.

Agarwal-Gardner Typecurve Model utilizes the concepts of equivalence between constant rate and constant pressure solutions. The dimensionless variables of the typecurves are based on conventional welltest definitions.	
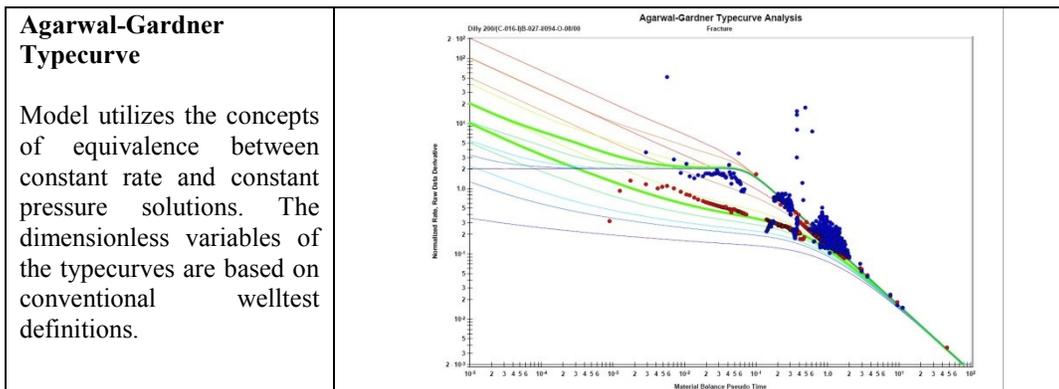
Blasingame Elliptical Flow Model assumes a hydraulically fractured well in the center of a reservoir with a closed elliptical boundary.	
Blasingame Finite Conductivity Fracture Model assumes a planar crack propagated from a well by hydraulic fracturing having a non-zero pressure drop in the fracture.	
Blasingame Infinite Conductivity Fracture Model assumes an infinite conductivity hydraulic fracture in the center of a cylindrical reservoir.	
Blasingame Horizontal Well Model assumes a square shaped reservoir with uniform thickness with a horizontal well penetrating the center of the pay zone.	

181

Normalized Pressure Integral Typecurve Model implements pressure integral curve as the base curve for diagnostics that does not suffer noisy and scattered data.	
Transient Finite Conductivity Fracture Typecurve Model assumes a square-shaped reservoir with a hydraulic fracture in the center.	
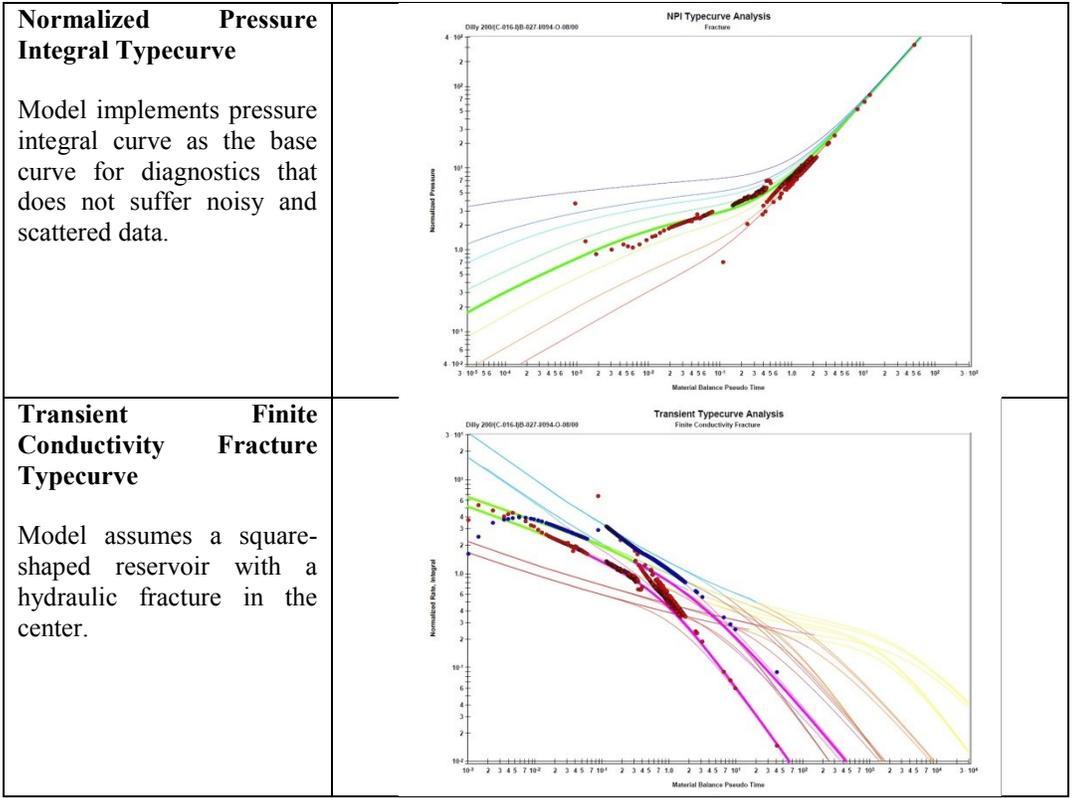
Figure 5-9 Summary of the model assumptions and type curve analysis for selected production data analysis techniques. Values of estimated parameters are shown in Table 5-3.

Chapter 6: Conclusions and Recommendations

This thesis has described novel re-sampling procedures to honor geologic information in reservoirs with non-Gaussian model parameters and a new workflow for characterization of fracture networks. The research methodology consists of generating multiple geological models and updating the uncertain parameters using dynamic flow responses using iterative EnKF technique.

Implementation of EnKF together with re-sampling of the new realizations demonstrates reasonable improvement in the history matching and uncertainty assessment of non-Gaussian and unconventional reservoir models. This chapter summarizes the main conclusions drawn from this work.

6.1 Contributions and Conclusions

6.1.1 Characterization of Facies distribution using ensemble Kalman filter with re-sampling

Two novel re-sampling procedures are proposed to honor geologic information in reservoirs with non-Gaussian model parameters after history matching. The methods entail combining a re-sampling step with the discrete cosine transform – ensemble Kalman filter approach.

Method 1.   The re-sampling step consists of constructing a facies probability map and application of probability field (P-Field) simulation to re-sample a new ensemble. After certain number of assimilation steps in EnKF, the updated ensemble members are used to propose a probability map for facies distribution. P-Field simulation is performed subsequently using

183

the facies probability map to generate a new ensemble, which honors the static geologic data and is more consistent with the early production data.

Method 2.   A novel probability weighted re-sampling scheme is implemented to generate a new ensemble using the updated ensemble members. After certain number of assimilation steps, model parameters having the highest impact on the production performance are selected as re-sampling points. These points capture updating information derived from previous assimilation steps to be incorporated in the new ensemble. The new ensemble is generated using the information at the re-sampling points as well as the reference statistics regarding the proportions and spatial continuity of different facies.

After resampling, the entire ensemble is subject to the forward model from the beginning until the last EnKF update step. This will certainly incur additional computational efforts. However, the additional costs can be justified by the improvement in terms of improving the ensemble variance and maintaining the ensemble diversity as well as reference statistics reproduction and uncertainty estimation in the posterior probability distribution.

6.1.2 History matching and uncertainty quantification of discrete fracture network models

An integrated approach for history matching and characterization of reservoirs with natural or hydraulic fractures is presented. This new methodology includes generating multiple discrete fracture network (DFN) models, upscaling the

models for numerical multiphase flow simulation, and updating the fracture parameters using dynamic flow responses such as continuous rate and pressure measurements.

Fracture parameters having the highest uncertainty and impact on the production performance are selected as optimization parameters. Subsequent to each optimization step, a new ensemble of DFN models are generated using updated parameters. The method used for dynamic data integration entails combining probability weighted sampling with Ensemble Kalman filter. The weighting scheme is based on the dynamic data mismatch. Implementation of this automated history matching approach results in multiple equally probable discrete fracture network models and their upscaled flow simulation models which honor the geological information and at the same time they match the dynamic production history.

The methodology is implemented for history matching of two different case studies. The first one is a synthetic, naturally occurring tight gas fractured reservoir model in which dynamic observed data, including gas production rate and well bottom hole pressure measurements, are integrated for characterization and uncertainty assessment of spatial distribution and orientation of the natural fractures in the reservoir. In the next application, the methodology is applied to facilitate characterization of hydraulic fracture parameters of a multi-stage hydraulic fractured shale gas well in the Horn River basin.

In this research Oda's techniques is used to upscale discrete fracture network models. Oda is a static approach without any flow simulation computations. This

method is fast and efficient, and it is well suited for well-connected and high-density fractures. For low density fracture sets, such as sparse distributed natural fractures in shale reservoirs, using Oda technique may degrade upscaling results and flow based upscaling methods are more accurate. However, computational efficiency and simulation costs should be considered when applying flow based upscaling approaches.

6.2 Recommendations for Future Work

For further improvement in dynamic data integration and characterization of non-Gaussian reservoirs, the following future research is recommended:

- The presented approaches to probability re-sampling of the ensemble are applied for characterization of two dimensional synthetic reservoir models. The approach could benefit from further study and application for model parameter estimation of real reservoir models such as SAGD reservoirs. For such applications relative permeability and capillary pressure of different flow units (facies) should be included in the state vector and updated together with permeability, porosity and facies indicators.

- The extra forward model execution after re-sampling incurs additional computational effort. The improvement in estimation accuracy of real reservoir models shall be investigated to justify the extra simulation costs.

- The hydraulic fracturing operation data such as injection pressure, total injected volume, injection rate and etc. can be used to simulate hydraulic

186

fracturing process and obtain invaluable knowledge of fracture pattern and network. These results improve the initial knowledge of the fracture parameters and their spatial distribution. Similarly, fluid production during flow back period in tight/shale gas reservoirs provides valuable information regarding hydraulic fracture parameters such as fracture volume, fracture surface area as well as rock properties such as capillary pressure curves. The fracture parameters derived from aforementioned simulations shall be integrated to generate an improved set of initial discrete fracture network models for history matching.

- Uncertainty assessment and history matching of natural and hydraulic fracture reservoirs using a two stage algorithm will significantly improve the overall performance of the methodology described in chapters four and five. The algorithm shall include a fast proxy approximation as an initial stage to analyze the rough sensitivities of the production performance to the fracture parameters. In the second stage, the discrete fracture network models that pass the proxy approximation stage will be subjected to upscaling and dynamic simulation.

References

Aanonsen, S.I., Nævdal, G., Oliver, D.S., Reynolds, A.C., Valles, B. (2009). The ensemble Kalman filter in reservoir engineering – a Review. SPE Journal, 14(3), 393-412.

Agarwal, R.G., Gardner, D.C., Kleinsteiber, S.W., Fussell, D.D. (1999). Analyzing well production data using combined type-curve and decline-curve analysis concepts. SPE Reservoir Evaluation and Engineering, 2(5):478-486.

Agbalaka, C., Oliver, D.S. (2008). Application of the EnKF and localization to automatic history matching of facies distribution and production data. Mathematical Geosciences, 40(4):353-374.

Al-Kobaisi, M., Ozkan, E., Kazemi, H. (2006). A hybrid numerical-analytical model of finite-conductivity vertical fractures intercepted by a horizontal well. SPE Reservoir Evaluation and Engineering, 9(4):345-355

Alkouh, A., Patel, K., Schechter, D., Wattenbarger, R. (2012). Practical use of simulators for characterization of shale reservoirs (SPE 162645). In: Proceedings of SPE Canadian Unconventional Resources Conference.

Amini, S., Ilk, D., Blasingame, T.A. (2007). Evaluation of the elliptical flow period for hydraulically-fractured wells in tight gas sands - Theoretical aspects and practical considerations (SPE 106308). In: Proceedings of SPE Hydraulic Fracturing Technology Conference.

Anderson, D.M., Stotts, G.W.J., Mattar, L., Ilk, D., Blasingame, T.A. (2006).

Production data analysis - Challenges, Pitfalls, Diagnostics (SPE 102048). In: Proceedings of SPE Annual Technical Conference and Exhibition.

Anderson, J.L. (2001). An ensemble adjustment Kalman filter for data assimilation. Monthly Weather Review, 129: 2884-2903.

Arroyo, E., Devegowda, D., Datta-Gupta, A., Choe, J. (2008). Streamline-assisted ensemble Kalman filter for rapid and continuous reservoir model updating. SPE Reservoir Evaluation and Engineering, 11(6): 1046-1060.

Barree, R.D., Cox, S.A., Gilbert, J.V., Dobson, M. (2005). Closing the gap: fracture half-length from design, buildup and production analysis. SPE Production and Facilities, 20(4):274–285

Bello, R.O/ (2009). Rate transient analysis in shale gas reservoirs with transient linear behavior. PhD Dissertation, Texas A &M University.

Blasingame, T.A., Johnston, J.L., Lee, W.J. (1989). Type-curve analysis using the pressure integral method (SPE 18799). In: Proceedings of SPE California regional meeting.

Bogdanov, I.I., Mourzenko, V.V., Thovert, J.F., Adler, P.M. (2007). Effective permeability of fractured porous media with power-law distribution of fracture sizes. Physical Reviews E, 76(036309) 1-15.

Boisvert, J.B., Pyrcz, M.J., Deutsch, C.V. (2008). Multiple-point statistics for training image selection. Natural Resources Research, 16(4): 313-321.

Bourbiaux, B., Cacas, M.C., Sarda, S., Sabathier, J.C. (1998). A Rapid and Efficient Methodology to Convert Fractured Reservoir Images Into a Dual-Porosity Model. Oil & Gas Science and Technology, 53(6) 785-799.

Brown, M., Ozkan, E., Raghavan, R., Kazemi, H. (2011). Practical solutions for pressure transient responses of fractured horizontal wells in unconventional shale reservoirs. SPE Reservoir Evaluation and Engineering 14(6):663–676

Burgers, G., van Leeuwen, P.J., Evensen, G. (1998). Analysis scheme in the ensemble Kalman filter. Monthly Weather Review, 126: 1719-1724.

Chang, H., Zhang, D., Lu, Z. (2010). History matching of facies distribution with the EnKF and level set parameterization. Journal of Computational Physics 229: 8011-8030.

Chen, Y., Durlofsky, L. J., Gerristen, M., Wen, X. H. (2003). A Coupled Local-Global Upscaling Approach for Simulating Flow in Highly Heterogeneous Formations. Advances in Water Resources, 26(10) 1041-1060.

Chen, Y., Oliver, D.S. (2010). Cross-covariances and localization for EnKF in multiphase flow data assimilation. Computational Geosciences, 14(4): 579-601.

Chen, Y., Oliver, D.S. (2012). Ensemble randomized maximum likelihood method as an iterative ensemble smoother. Mathematical Geosciences,

44(1): 1-26.

Chen, Y., Zhang, D. (2006). Data assimilation for transient flow in geologic formations via ensemble Kalman filter. Advances in Water Resources, 29(8):1107–1122.

Cheng, Y., Lee, W.J., McVay, D.A. (2009). A new approach for reliable estimation of hydraulic fracture properties using elliptical flow data in tight gas wells. SPE Reservoir Evaluation and Engineering, 12(2):254–262.

Clarkson, C.R. (2011). Integration of Rate-Transient and Microseismic Analysis for Unconventional Gas Reservoirs: Where Reservoir Engineering Meets Geophysics. Recorder, Canadian Society of Exploration Geophysicists, 36(10):44–61.

Clarkson, C.R. (2013a). Production data analysis of unconventional gas wells: Review of theory and best practices. International Journal of Coal Geology, 109–110:101–146.

Clarkson, C.R. (2013b). Production data analysis of unconventional gas wells: Workflow. International Journal of Coal Geology, 109–110:147–157.

Cliffe, K., Holton, D., Houston, P., Jackson, C., Joyce, S., Milne, A. (2011). Conditioning discrete fracture network models of groundwater flow. International Journal of Numerical Analysis and Modeling, 8(4) 543-565.

Computer Modeling Group (2013). GEM compositional and unconventional reservoir simulator: user's guide, Computer Modeling Group LTD.

http://www.cmgl.ca.

Cressie, N. (1985). Fitting variogram models by weighted least squares. Mathematical Geology, 17(5): 563-586.

Cui, H., Kelkar, M. (2005). Automatic History Matching of Naturally Fractured Reservoirs and a Case Study (SPE 94037). In: Proceedings of SPE Western Regional Meeting.

De Lima, A., Lange, A., Schiozer, D.J. (2012). Assisted History-Matching for the Characterization and Recovery Optimization of Fractured Reservoirs Using Connectivity Analysis (SPE 154392). In: Proceedings of SPE Europec/EAGE Annual Conference.

de Marsily, G., (1984). Spatial Variability of Properties in Porous Media: A Stochastic Approach, in Fundamentals of Transport in Porous Media. Proceedings of NATO Advanced Study Institute on Mechanics of Fluids in Porous Media, Bear J. and Corapcioglu M. Y., Martinus Nijhoff, Boston. 719-769.

Dershowitz, W.S., (1993). Geometric Conceptual Models for Fractured Rock Masses: Implications for Groundwater Flow and Rock Deformation. Proceedings, Eurock-93, Second European Conference of the International Society for Rock Mechanics, Lisbon.

Dershowitz, W.S., P. La Pointe, T. Doe, (2004). Advances in Discrete Fracture Network Modeling. In Proceedings of the US EPA/NGWA Fractured Rock Conference, 2004 - info.ngwa.org.

Deutsch, C.V., Journel, A.G. (1997). GSLIB: Geostatistical Software Library and User's Guide. Second edition, Oxford University Press, New York.

Dovera, L., Della Rossa, E. (2010). Multimodal Ensemble Kalman Filtering Using Gaussian Mixture Models. Computational Geosciences, 15(2), 307-323.

Emerick, A.A., Reynolds, A.C. (2011). Combining sensitivities and prior information for covariance localization in the ensemble Kalman filter for petroleum reservoir applications. Computational Geosciences, 15(2): 251-269.

Emerick, A.A., Reynolds, A.C. (2012). Combining the ensemble Kalman filter with Markov Chain Monte Carlo for improved history matching and uncertainty characterization. SPE Journal, 17(2): 418-440.

Emerick, A.A., Reynolds, A.C. (2013). Investigation of the sampling performance of ensemble-based methods with a simple reservoir model. Computational Geosciences, 17(2): 325-350.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research, 99(C5): 10143-10162.

Evensen, G. (2003). The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dynamics 53: 343-367.

Evensen, G. (2007). Data Assimilation: The Ensemble Kalman Filter. Berlin: Springer.

Fekete (2014) Theory and equations, Fekete Associates. http://www.fekete.com.

Fisher, M.K., Wright, C.A., Davidson, B.M., Steinsberger, N.P., Buckler, W.S., Goodwin, A., Fielder, E.O. (2005). Integrating fracture mapping technologies to improve stimulations in the Barnett shale. SPE Production Facilities, 20(2):85–93.

Fisher, R., (1953). Dispersion on a sphere. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences. 217(1130) 295-305.

FRACMAN 7 (2011) User Documentation, Golder Associates.

Froidevaux, R. (1993). Probability field simulation. In: A. Soares (ed) Geostatistics Troia '92, Dordrecht, The Netherlands, Kluwer Academic Publishers, 2, 73–84.

Fung, L.S.K., Collins, D.A., (1991). An Evaluation of the Improved Dual Porosity Model for the Simulation of Gravity Effects in Naturally Fractured Reservoirs. Journal of Canadian Petroleum Technology, 30(3) 61-68.

Furrer, R., Bengtsson, T. (2007). Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants. Journal of Multivariate Analysis, 98(2): 227-255.

Gang, T., Kelkar, M. (2008). History Matching for Determination of Fracture Permeability and Capillary Pressure. SPE Reservoir Engineering, 11(5) 813-822.

Ghods, P., Zhang, D. (2010). Ensemble Based Characterization and History Matching of Naturally Fractured Tight/Shale Gas Reservoirs (SPE 133606). In: Proceedings of SPE Western Regional Meeting.

Ghods, P., Zhang, D. (2012). Automatic Estimation of Fracture Properties in Multi-Stage Fractured Shale Gas Horizontal Wells for Reservoir Modeling (SPE 153913). In: Proceedings of SPE Western Regional Meeting.

Gomez, S., Gosselin, O., Barker, J.W. (2001). Gradient-based History- Matching with a global optimization method. SPE Journal, 6(2), 200-208.

Goodman, R. (1999). Introduction to Stochastic Models, the Benjamin / Cummins Publishing Company.

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

Gul, A., Nejadi, S., Shah, S., Trivedi, J. (2011). Make use of dynamic data – A constrain based EnKF for SAGD reservoir characterization and production management. Paper WHOC 11-568 presented at World Heavy Oil Congress held in Edmonton, Alberta, Canada, March 14-17.

Houtekamer, P.L., Mitchell, H.L. (2000). An adaptive ensemble Kalman filter. Monthly Weather Review, 128: 416-433.

Hu, L.Y., Jenni, S. (2005). History Matching of Object-Based Stochastic Reservoir Models. SPE Journal, 10(3) 312-323.

Jafarpour, B., McLaughlin, D.B. (2008). History matching with an ensemble Kalman filter and discrete cosine parameterization. Computational Geosciences, 12(2): 227-244.

Jafarpour, B., Tarrahi, M. (2011). Assessing the Performance of the Ensemble Kalman Filter for Subsurface Flow Data Integration under Variogram Uncertainty. Water Resources Research, 47(5): W05537.

Jain, A.K. (1988). Fundamentals of Digital Image Processing. Prentice-Hall, Englewood Cliffs, NJ.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. Journal of Basic Engineering. 82(1): 35–45.

Karami-Fard, M., Gong, B., Durlofsky,L. J. (2006). Generation of coarse-scale continuum flow models from detailed fracture characterizations. Water Resources Research, 42(W10423) 1-13.

Kazemi, H., Merrill, L.S., Porterfield, K.L, Zeman, P.R. (1967). Numerical Simulation of Water-Oil Flow in Naturally Fractured Reservoirs. SPE Journal, 16(6):317-326.

Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P. (1983). Optimization by simulated annealing. Science 220 (4598): 671–680.

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

Li, L., Zhou, H., Hendricks Franssen, H. J., Gómez-Hernández, J. J. (2012).

Groundwater flow inverse modeling in non-MultiGaussian media: performance assessment of the normal-score Ensemble Kalman Filter. Hydrology and Earth System Sciences, 16(2): 573-590.

Li, L., Zhou, H., Hendricks Franssen, H.J., Gómez-Hernández, J.J. (2012). Pattern recognition in a bimodal aquifer using the normal-score ensemble Kalman filter. Mathematical Geosciences, 44(2):169-185.

Liu, N., Oliver, D.S. (2005). Ensemble Kalman Filter for Automatic History Matching of Geologic Facies. Journal of Petroleum Science and Engineering, 47(3-4): 147-161.

Lorentzen, R.J., Flornes, K.M., Nævdal, G. (2012). History matching channelized reservoirs using the ensemble Kalman filter. SPE Journal, 17(1): 137-151.

Lorentzen, R.J., Fjelde, K.K., Froyen, J., Lage, A.C., Naevdal, G., Vefring, E.H. (2001). Underbalanced and Low-Head Drilling Operations: Real Time Interpretation of Measured Data and Operational Support (SPE 71384). In: Proceedings of SPE Annual Technical Conference and Exhibition.

Luo, H., Mahiya, G., Pannett, S., Benham, P.H. (2014). The use of rate-transient-analysis modeling to quantify uncertainties in commingled tight gas production-forecasting and decline-analysis parameters in the Alberta Deep basin. SPE Reservoir Evaluation and Engineering 17(2):209–219.

Ma, X., Al-Harbi, M., Datta-Gupta, A., Efendiev, Y. (2008). An efficient two-stage sampling method for uncertainty quantification in history matching

geological models. SPE Journal, 13(1):77-87.

Medeiros, F., Ozkan, E., Kazemi, H. (2008). Productivity and drainage area of fractured horizontal wells in tight gas reservoirs. SPE Reservoir Evaluation and Engineering, 11(5):902–911.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. The Journal of Chemical Physics, 21(6): 1087.

Modis, K., Sideri, D. (2013). Geostatistical Simulation of Hydrofacies Heterogeneity of the West Thessaly Aquifer Systems in Greece. Natural Resources Research. 22(2), 123-138.

Moghadam, S., Jeje, O., Mattar, L. (2011). Advanced gas material balance in simplified format. Journal of Canadian Petroleum Technology, 50(1):90–98.

Mood, A.M. (1940). The distribution of runs. The Annals of Mathematical Statistics, 11(4) 367-392.

Moreno, D., Aanonsen, S.I. (2011). Continuous Facies Updating Using the Ensemble Kalman Filter and the Level Set Method. Mathematical Geosciences, 43(8): 951-970.

Myrseth, I., Sætrom, J., More, H. (2012). Resampling the ensemble Kalman filter. Computational Geosciences 55: 44-53.

Nash, J.E., Sutcliffe, J.V. (1970). River flow forecasting through conceptual

models part I - A discussion of principles. Journal of Hydrology, 10(3): 282-290.

Nejadi, S., Trivedi, J., Leung, J. (2011). Improving characterization and history matching using entropy weighted ensemble Kalman filter for non-Gaussian distributions (SPE 144578). In: proceedings of SPE Western North American Regional Meeting.

Nejadi, S., Leung, J., Trivedi, J. (2012a). Integration of production data for estimation of natural fracture properties in tight gas reservoirs using ensemble Kalman filter (SPE 162783). In: proceedings of Canadian Unconventional Resources Conference.

Nejadi, S., Trivedi, J., Leung, J. (2012b). Ensemble Kalman filter predictor bias correction method for non-Gaussian geological facies detection. In: proceedings of IFAC Conference on Automatic Control in Offshore Oil and Gas Production (ACOOG 2012).

Nejadi, S., Trivedi, J., Leung, J. (2012c). Estimation of geological facies boundaries using categorical indicators with P-Field simulation and ensemble Kalman filter (SPE 153645). In: Proceedings of SPE Latin America and Caribbean Petroleum Engineering Conference.

Nejadi, S., Leung, J., Trivedi, J. (2014a). Characterization of non-Gaussian geologic facies distribution using ensemble Kalman filter with probability weighted re-sampling. Mathematical Geosciences, doi:10.1007/s11004-014-9548-8.

Nejadi, S., Trivedi, J., Leung, J. (2014b). Estimation of geological facies boundaries using categorical indicators with P-Field simulation and ensemble Kalman filter. Natural Resources Research, doi:10.1007/s11053-014-9233-0.

Oda, M. (1985). Permeability tensor for discontinuous rock masses. Geotechnique, 35(4) 483-495.

Oliver D.S., Cunha, L.B., Reynolds, A.C. (1996). Markov chain Monte Carlo methods for conditioning a permeability field to pressure data. Mathematical Geology, 29(1): 61-91.

Oliver, D.S., Chen, Y. (2011). Recent progress on reservoir history matching: a review. Computational Geosciences, 15(1): 185-221.

Oliver, D.S., Reynolds A.C., Liu, N. (2008). Inverse theory for petroleum reservoir characterization and history matching. Cambridge University Press, Cambridge.

Ouenes, A. (2000). Practical application of fuzzy logic and neural networks to fractured reservoir characterization. Computers and Geosciences, 26(9): 953-962.

Ouenes, A., Richardson, S., Weiss W.W. (1995). Fractured Reservoir Characterization and Performance Forecasting Using Geomechanics and Artificial Intelligence (SPE 30572). In: Proceedings of SPE Annual Technical Conference and Exhibition.

Ping, J., Zhang, D. (2013). History matching of fracture distributions by

ensemble Kalman filter combined with vector based level set parameterization. Journal of Petroleum Science and Engineering, 108: 288-303.

Pyrcz, M.J., Deutsch, C.V. (2001). Two Artifacts of Probability Field Simulation. Mathematical Geology, 33(7): 775-799.

RamaRao, B.S., LaVenue, A.M., de Marsily, G., Marietta, M.G. (1995). Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields 1. Theory and computational experiments. Water Resources Research, 31(3): 475-493.

Rao, K.R., Yip, P. (1990). Discrete Cosine Transform: Algorithms, Advantages, Applications. Academic Press, Boston.

Remy, N., Boucher, A., Wu, J. (2009). Applied Geostatistics with SGeMS: A User's Guide, Cambridge University Press, Cambridge.

Sahni, I., Horne R.N. (2005). Multiresolution wavelet analysis for improved reservoir description. SPE Reservoir Engineering and Evaluation, 8(1):53-69.

Saleri, N.G., Toronyi, R.M., Shyder, D.E. (1992). Data and data hierarchy. Journal of Petroleum Technology, 44(12): 1286-1293.

Sarma, P., Durlofsky, L.J., Aziz, K. (2008). Kernel principal component analysis for efficient, differentiable parameterization of multipoint geostatistics. Mathematical Geosciences, 40(1): 3-32

Sarma, P., Durlofsky, L.J., Aziz, K., Chen, W.H. (2007). A new approach to automatic history matching using Kernel PCA (SPE 106176). In: proceedings of SPE Reservoir Simulation Symposium.

Schlumberger Information Solutions (2011) Eclipse reservoir simulator: Reference manual and technical description. http://www.slb.com.

Srinivasan, S., Cares, J. (2000). Conditioning reservoir models to dynamic data – A forward modeling perspective (SPE 62941). In: Proceedings of SPE Annual Technical Conference and Exhibition.

Srivastava, R.M. (1992). Reservoir characterization with probability field simulation. SPE Formation Evaluation, October, 927-937.

Suzuki, S., Daly, C., Caers, J.K., Mueller, D. (2007). History Matching of Naturally fractured Reservoirs Using Elastic Stress Simulation and Probability Perturbation Method. SPE Journal, 12(1): 118-129.

Teng, Y., Koike, K. (2007). Three-dimensional imaging of a geothermal system using temperature and geological models derived from a well-log dataset. Geothermics, 36(6): 518-538.

van Leeuwen, P.J. (2003). A variance-minimizing filter for large-scale applications. Monthly Weather Review, 131: 2071-2084

Warren, J.E., Root, P.J. (1963). The Behavior of Naturally Fractured Reservoirs. SPE Journal, 3(3): 245-255.

Wattenbarger, R.A., El-Banbi, A.H., Villegas, M.E., Maggard, J.B. (1998).

Production analysis of linear flow into fractured tight gas wells (SPE39931). In: Proceedings of SPE Rocky Mountain regional/low-permeability reservoirs symposium and exhibition.

Yamamoto, J.K., Mao, X.M., Koike, K., Crosta, A.P., Landim, P.M.B., Hu, H.Z., Wang, C.Y., Yao, L.Q. (2012). Mapping an uncertainty zone between interpolated types of a categorical variable. Computers and Geosciences, 40: 146-152.

Zafari, M., Reynolds, A.C. (2007). Assessing the Uncertainty in Reservoir Description and Performance Predictions with the Ensemble Kalman Filter. SPE Journal, 12(3): 382-391.