PRESERVATION AND IDENTIFICATION OF LARGE SCALE CONNECTIVITY FROM PRODUCTION DATA

A Dissertation

by

MORTEZA KHODABAKHSHI

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

DOCTOR OF PHILOSOPHY

Chair of Committee,	Behnam Jafarpour
Co-Chair of Committee,	Michael J. King
Committee Members,	Eduardo Gildin
	Faming Liang
Head of Department,	A. Daniel Hill

May 2014

Major Subject: Petroleum Engineering

ABSTRACT

Multipoint statistics (MPS) provides an approach for pattern-based simulation of complex geologic objects from a training image (TI), which contains the general connectivity structures of complex patterns. While grid-based implementation of the MPS methods facilitates hard-data conditioning, conditioning the simulated facies on flow data poses a challenging problem. The main objective of this dissertation is to develop an inverse modeling framework for conditioning MPS-based facies simulation on dynamic flow data. The developed formulation is then extended to account for uncertainty in the geologic scenario. In the second part of the dissertation, an inverse modeling formulation is presented for estimating large-scale reservoir connectivity from low-resolution pressure data using.

The first contribution of this dissertation is the formulation of a probability conditioning method (PCM). In the PCM approach, the flow data is first inverted to obtain a probabilistic description of facies distribution (a probability map). The resulting probability map is then used to guide the MPS facies simulation from a specified TI. The proposed PCM approach can be used with different inversion algorithms. In this dissertation two alternative implementations are presented: 1) the ensemble Kalman filter (EnKF); 2) a stochastic optimization approach. An important practical limitation of the MPS modeling approach is the uncertainty in the prior TI. To address this problem, a Bayesian mixture modeling formulation is developed. In this approach, the posterior distribution of facies is partitioned into individual conditional densities of the TIs.

In the second part of the dissertation, a novel approach is developed estimating field-scale reservoir connectivity from dynamic data. This is accomplished by reconciling low-resolution dynamic field pressure data with high-resolution static geologic models. Since pressure variation represents a smooth function, a low-resolution (coarse-scale) grid system is adopted for reservoir simulation. To reconcile data and model resolutions, the grid system is generated using the Delaunay triangulation method. The reservoir properties are then scaled up from the fine-scale geological model to this coarse scale unstructured grid system to create an initial static simulation model. In the data integration stage, the EnKF is used to automatically adjust the global parameters of the field to match the static pressure.

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

Page

1. INTRODUCTION

1.1. History Matching Problem

Subsurface systems pose some of the most challenging characterization and modeling problems in science with significant environmental, public health, and energy security implications. The main difficulties in understanding and modeling subsurface phenomena are related to inaccessibility and heterogeneity of geologic formations as well as complex interactions between fluids and rocks over a wide range of temporal and spatial scales. Consequently, significant uncertainty is introduced into predictions of the related flow and transport processes, thereby complicating the development of subsurface hydrological, energy, mineral, and environmental resources.

The progress in numerical modeling of subsurface flow and transport processes within the past 50 years has played a significant role in managing hydrocarbon resources (Peaceman and Rachford 1955; Aziz and Settari 1979). While forward flow simulation is critical for prediction of multiphase fluid flow and transport behavior in geologic formations, the quality of such predictions ultimately depends on the accuracy and representativeness of the models used to generate them. Therefore, parallel to numerical forward modeling, integration of complex geologic models with diverse datasets using inverse modeling has been applied to calibrate numerical models of hydrocarbon reservoirs (e.g., Cooley 1979; Yeh and Yoon 1981; Hill and Tiedeman 2007; Oliver et al. 2008; Carrera et al. 2005; Hendricks-Franssen et al. 2009).

A particularly important aspect of inverse modeling is uncertainty quantification (e.g., Moore and Doherty 2005). Simplifying assumptions in forward modeling of physical processes in subsurface systems, together with inaccurate representations of heterogeneity in complex geologic formations and structural and parameter heterogeneity lead to significant uncertainty in developing underground water and energy resources (Blazkova and Beven, 2009; Gotzinger and Bardossy, 2008; Solomatine and Shrestha, 2009; Thyer et al., 2009; Tonkin and Doherty, 2009; Zhang et al., 2008).

The main source of uncertainty in subsurface modeling and characterization is data scarcity which results from the high costs and technical difficulties in accessing and sampling from deep underground formations. Advances in data acquisition systems and remote sensing technologies are presenting promising prospects for alleviating data limitation issues (Yeh et al., 2008). At the same time, rapid growth in computational power and improved understanding of the dynamics of subsurface flow and transport systems is facilitating the construction of more complex and sophisticated predictive models. Progress in these areas is expected to lead to more accurate and better informed modeling exercises. However, building complex models without proper parameterization and rigorous calibration against pertinent sources of data with different scales, accuracy, and coverage can result in inaccurate representation of reality. In particular, failure to adequately account for the uncertainty in the spatial distribution of flow and transport related properties can lead to unreliable predictions of fluid displacement patterns, which adversely influence field development strategies.

The dynamic response of a reservoir to forced disturbances contains valuable information pertaining to both local trends in fluid flow properties as well as global geologic facies connectivity in the reservoir (Alcolea et al. 2010; Renard and Caers 2008). Constraining subsurface flow models to dynamic measurements of pressure or phase flow rates is considerably more complicated than integration of static data. The nonlinear mapping of input model parameters (model space) onto dynamic reservoir responses (data space), and the integral (spatially averaged) and sparse nature of the available dynamic data are among the main contributors to the complexity of the resulting model calibration inverse problem. Over the last decade, considerable efforts have been devoted to developing numerical inverse modeling approaches for calibration of reservoir models and quantification of the uncertainty in model descriptions and the resulting flow predictions (e.g., Sun, 1994; de Marsily et al., 1999; Carrera et al., 2005, Yeh et al., 2007; Hill and Tiedeman 2007; Oliver et al. 2008). Various deterministic and probabilistic inversion techniques have been developed and applied to subsurface flow-model calibration problems.

Deterministic inverse methods minimize a user-defined cost function that usually penalizes discrepancies between predicted and observed dynamic and static data, as well as departure from direct and/or indirect prior information about the solution. The literature clearly establishes that inference of heterogeneous hydraulic rock properties such as spatial distribution of permeability from flow measurements typically leads to an ill-posed nonlinear dynamic inverse problem with several nonunique solutions. That is, different parameter combinations can reproduce the observed (past) response of a

reservoir equally well, yet provide different future predictions (Yeh, 1986; Carrera and Neumann, 1986a-c; Carrera, 1987; McLaughlin and Townley, 1996; de Marsily et al., 1999; Carrera et al., 2005; Hill and Tiedeman, 2007; Oliver et al., 2008).

The solution of ill-posed inverse problems is also hampered by numerical instabilities in solution algorithms. Two common methods for mitigating instability and nonuniqueness issues associated with the solution of ill-posed inverse problems are parameterization and regularization. Parameterization methods usually aim at reducing the number of unknown model parameters to formulate and solve a better posed inverse problem (e.g., Jacquard and Jain 1965, Jahns 1966; Shah et al. 1978; de Marsily et al. 1984;Carrera and Neuman 1986a-c; Doherty 2003; Gavalas et al. 1976; Reynolds et al., 1996; Jafarpour and McLaughlin 2009a; dos Santos Amorim et al., 2010). Regularization refers to incorporation of direct or indirect prior information into the solution algorithm, often in the form of constraints (e.g., Tikhonov and Arsenin 1977; Lawson and Hanson 1995; Weiss and Smith 1998; Doherty and Skahill 2006; Linden et al. 2005; Tonkin and Doherty 2005; Carrera and Neuman, 1986a-c; Moore and Doherty, 2005; Hill and Tiedeman 2007; Oliver et al., 2008).

Probabilistic methods, on the other hand, seek to characterize the solution of inverse problems in terms of a conditional probability distribution. The Bayesian inversion theory provides an elegant framework for characterization of the probability distribution of model parameters after conditioning on measurements. Variants of the Bayesian inversion aim at integrating a prior probability density function (PDF) with a

likelihood model that describes the probability of producing the data given a set of model parameters (Tarantola, 2004).

Analytical solutions to the Bayesian inversion framework are limited to a very few special cases where, for example, the prior and posterior distributions are conjugate; that is, the likelihood function is such that the prior and posterior PDFs belong to the same family of distributions. Therefore, the main practical applications of the Bayesian inversion to large-scale nonlinear inverse problems include: estimating point statistics of the conditional PDF, such as its central tendency (mean, mode, or median) and a measure of dispersion (covariance); and approximating the conditional distribution by estimating a finite number of samples from it.

The latter approach has become particularly popular in recent years, due primarily to availability of powerful and inexpensive computational resources and development of relatively simple ensemble model calibration techniques, making it possible to conduct systematic uncertainty quantification and risk assessment analysis (Sahuquillo et al., 1992; LaVenue et al., 1995; RamaRao et al., 1995; Gomez-Hernandez et al., 1997; Sambridge and Mosegaard, 2002; Lorentzen et al., 2003; Naevdal et al., 2005; Chen and Zhang, 2006;Wen and Chen, 2006; Nowak, 2009).

1.2. Preserving Large Scale Feature Continuity

Many of the existing inversion techniques are suitable for calibrating flow models with spatially variable parameters that can be captured by up to second-order (two-point) statistical characterization. While conventional second-order geostatistics is widely

applied to represent the variability in spatial distribution of hydraulic properties in subsurface reservoir models, the connectivity structures in many geologic formations are far too complex to be amenable to second-order descriptions. The popularity of variogram-based modeling techniques is rooted more in their mathematical simplicity, computational efficiency, and implementation ease than in their geological interpretation and realism.

While the spatial variability of hydraulic properties in certain geologic formations is amenable to second-order characterization, many complex geologic structures such as those containing discrete geologic objects with sharp discontinuities across facies boundaries cannot be described with two-point statistical techniques (e.g., Gomez-Hernandez and Wen, 1998; Deutsch and Journel 1998; Carle et al. 1998, Western et al. 2001; de Marsily et al. 2005; Zinn and Harvey 2003). Of particular importance in subsurface flow and transport are the extreme phenomena that induce preferential flow paths (such as channels) or flow barriers (such as thin shale layers) that can dominate the behavior of local and global flow regimes. These complex extreme features do not lend themselves to conventional second-order geostatistical modeling descriptions. In addition, stochastic processes with distinctly different probability distributions (or higher-order statistics) can sometimes be indistinguishable when only their second-order characterization is considered, indicating the importance of higher-order statistics in describing geologic formation with complex spatial connectivity, such as discrete geologic objects (Strebelle, 2002; Caers et al., 2002).

Two common approaches for generating multiple realizations of geologic facies that honor a prior statistical characterization and various types of measured and interpreted data are pixel-based approaches such as sequential indicator simulation (Journel,1983; Isaaks,1990; Srivastava, 1992; Goovaerts, 1997; Chiles and Delfiner, 1999), and object-based (Boolean) methods, e.g. marked point process, that are better able to describe the continuity in geologic objects with well-defined shapes (Haldorsen and Lake, 1984; Stoyan et al., 1987; Deutsch and Wang, 1996; Holden et al., 1998). Object-based methods lack the flexibility of grid-based simulation techniques. As a result, data integration becomes particularly cumbersome in an object-based modeling framework (Honarkhah et al., 2010).

Multiple-point statistics (MPS) simulation methods (Guardiano and Srivastava, 1993; Strebelle, 2002, Caers and Zhang, 2004) provide a pattern-imitating, flexible, grid-based geostatistical simulation approach with the ability to reproduce complex geological patterns that cannot be modeled with variogram-based methods. Instead of using merely point-to-point statistics represented with a variogram model, MPS imitates the higher-order statistical patterns from a prior geologic model represented by a training image (TI).

Because of its grid-based implementation, conditioning MPS realizations to hard (for example, from a well) and soft (such as 3D seismic) data is not difficult (Strebelle, 2002, Journel, 2002, Remy et al., 2009). Nonetheless, the complexity in spatial variability of geologic facies, together with insufficient static data to adequately constrain simulated facies, can lead to significant uncertainties in describing facies

distributions in large models. Although recent progress in MPS has made it possible to describe complex non-Gaussian spatial structures, calibrating the output of MPS simulation against dynamic flow data remains an important research area.

Dynamic data provide additional useful information that can be used to further constrain the spatial variability in MPS-simulated facies realizations. The nonlinear relation between hydraulic properties and dynamic flow data complicates the process of constraining MPS simulation results to reproduce flow measurements. Caers and Hoffman (2006) developed a probability perturbation method (PPM) that makes it possible to simulate facies realizations from a given TI, conditioned on nonlinear flow data. In the PPM approach, a parameterization is used to perturb the pre-posterior probabilities that are used to generate new facies realizations. During the history matching iterations, flow data conditioning is accomplished by generating a new facies realization and comparing its flow prediction performance with that obtained from the current (iteration) realization using a predefined distance function (such as a flow data misfit norm). Perturbation of the simulation probabilities is continued during the history matching iterations by changing a single perturbation parameter (the random seed). The process is repeated until the generated realization is able to reproduce the observed data (within a specified error threshold) or the maximum number of iteration is reached (Caers, 2002; Hoffman and Caers, 2005; Kashib and Srinivasan, 2006). At the end of the iterations, the facies realization with the smallest distance (data misfit) is selected as a conditional realization. One advantage of the PPM approach is that the solutions are constructed to be consistent with the prior information (the TI). However, since, in

generating new facies, the probability perturbations are carried out without any feedback information from the flow data, new (proposed) facies iterates from the TI are drawn without a guideline to increase the likelihood of their acceptance. In light of the high computational cost associated with flow simulations, implementing the sampling process in PPM can be very expensive. The efficiency of the method can be significantly improved if flow data is used to actively guide the sampling process and to speed up the convergence of the method.

Alcolea and Renard (2010) present a block moving window algorithm for conditioning MPS simulations based on the block Markov Chain Monte Carlo method and use their method for integrating the pressure and connectivity data (Fu and Gomez-Hernandez, 2009). Mariethoz et al. (2010) developed the iterative spatial resampling method as an alternative approach to guide the simulation of new realization to honor the observed data. Zhou et al. (2012) proposed a pattern searching inverse approach to infer static and dynamic state variables in an MPS approach using an ensemble of conductivity samples and their respective heads as training data.

A standing challenge in MPS simulation and model calibration, however, remains: the uncertainty in the prior TI used to generate the conditional samples. This issue becomes particularly important considering the strict pattern-imitating nature of MPS simulation that limits the spatial variability of the resulting facies to the structural connectivity and existing patterns in the given TI. **Fig. 1.1** shows a satellite view of a section of Mississippi River near Baton Rouge. The river structure, orientation, and thickness vary in different regions, implying that the distribution of naturally occurring

features, such as the fluvial systems, can be too complex to represent with a single stationary TI. As depicted in Fig 1.1, the consistent TI for the fluvial system inside the left box is different from that on the right, even though the two sections of the river are close to each other. Underground fluvial or turbidite systems portray a similar complexity.

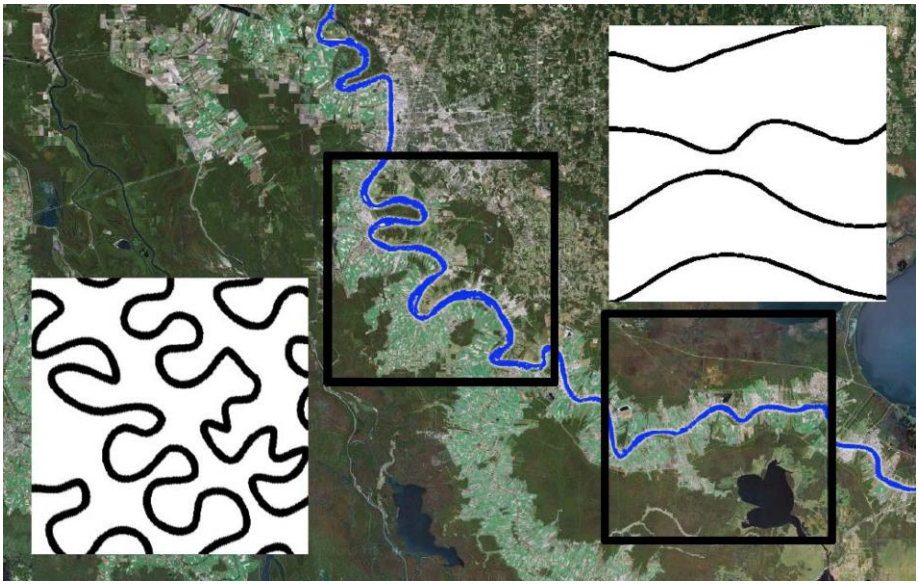


Fig. 1.1 – Sections of the Mississippi River as an example of naturally occurring fluvial systems with varying channel structure and orientation. Two sections of the river with (left) strong meandering features and (right) straight channels are highlighted. The corresponding conceptual models (TIs) are also shown next to each selected area.

Capturing discrete geological features is particularly important in predicting flow and transport processes because they strongly constrain fluid displacement patterns. Failure to correctly represent extreme geologic features can result in inaccurate predictions with serious field development and management consequences. One approach to deal with the uncertainty in describing the geologic continuity in a TI is to consider several TIs that capture the full range of geologic variability for a given formation. These TIs could be obtained based on different plausible geological scenarios, for example, from independent interpretations by different geologists or by stochastic treatment of parameters in a geologic modeling study that is used to identify possible connectivity patterns in the formation. Park et al (2013) rejected data-inconsistent training images using production data and then they used regional PPM for history matching.

In this dissertation, a novel approach was developed for conditioning MPS facies simulation on nonlinear dynamic flow data. The approach is general and can be implemented using various nonlinear inversion techniques. It uses the production data to infer the probability with which a particular facies type occurs at each grid block throughout the reservoir. It will then use this spatial probability map with the snesim algorithm to draw conditional facies realizations from a specified TI. The ensemble Kalman filter (EnKF), a recursive data assimilation technique, was used to integrate the flow data and infer facies probability maps.

Beside the depositional facies, other features like fractures and faults control the fluid flow behaviors of reservoirs. In contrast to the depositional facies, which has static

properties over a long period of time, the properties of these features are changed with time more frequently. For example, open fractures in one direction could close because of the changes in the stress-strain field in the reservoir, and new set of fractures may open to flow. The rates obtained from injection and production wells may be used to determine interwell connectivities.

The importance of interwell connectivity has prompted several studies in the literature. Heffer et al. (1997) estimated the connectivity of well pairs by using the Spearman rank correlation between the flow rates of injectors and producers. Refunjol and Lake (1999) used the same approach and maximized the rank correlation by considering a time lag that compensated for the effect of reservoir fluid compressibility. Soeriawinata and Kelkar (1999) developed an algorithm to include the superposition effect of multiple injectors on a producer. Panda and Chopra (1998) used an artificial neural network to identify interwell connectivity of injectors and producers. Demiroren (2007) transformed the injection/production measurements to the frequency domain to investigate the interwell connectivities. To analyze the injection/production measurements, Albertoni and Lake (2003) developed a model based on multivariate linear regression. In this model, the injection and total production rates are treated as input and output signals, respectively. They used this approach to predict the future production rates. Jensen et al. (2004) used the number of streamlines from an injection well to a production well to define a connectivity parameter. Dinh and Tiab (2008) used BHP data, measured under multi well pressure test similar to interference test, as the input and output signals. Yousef et al. (2006) developed the capacitance model (CM) for

estimating interwell connectivity, which was later applied to field data to estimate the permeability trend. The CM has also been used to estimate oil production rate and optimize water injection rate in a reservoir (Liang et al., 2007; Sayarpour et al., 2007). Webber et al. (2009) have provided some guidelines for applying the CM to large-scale field data.

While very simple and practical, the CM is derived based on simplifying assumptions (Kaviani, 2009; Yousef, 2006; Yousef et al., 2006) which are:

- The number of active producers must be constant during the time frame.
- BHP must be known or assumed constant.
- The productivity index of producers should be constant.
- The rate and pressure measurement should be accurate.
- The reservoir should be close to unit-mobility displacement and contain slightly compressible fluids.

Violation of the above assumptions may lead to failure of the CM approach to estimate the connectivities. Kaviani et al. (2012) have provided segmented and compensated CMs to relax some of the limiting assumptions in the CM. An outstanding limitation of the CM include lack of geologic description and spatial information about reservoir properties. In addition, the requirement to have injection wells in the field limits the application of the method under primary depletion during early development stages or production under aquifer pressure support.

Early detection of global reservoir connectivity is critical for improving field development strategies. Since the static pressure is not affected by local properties

around the well, static pressure, obtained from drilled or shut-in wells, provides valuable and useful information about large-scale reservoir continuity/discontinuity and strength of external sources like the aquifer. Analysis of this pressure data may be used to improve reservoir characterization and management, optimize oil recovery, and identify reservoir compartments.

In the next sections, overview of MPS facies simulation and EnKF data assimilation approaches is presented.

1.3. Multipoint Statistics

While the proposed methods are general approaches that can be used with any TI-based simulation method (such as snesim (Strebelle 2002), filtersim (Zhang et al. 2006), and simpat (Arpat and Caers, 2007) algorithms), the application of approach with the snesim algorithm is disscussed. The snesim facies simulation algorithm is suitable for modeling geologic formations that contain predominantly discrete objects such as fluvial and turbidite channels (Strebelle 2002; Caers and Zhang, 2004). In the snesim algorithm, a stationary TI represents the conceptual geologic continuity model in a given formation and is used to simulate multiple facies realizations. This pattern-imitating simulation framework preserves the higher-order spatial statistics in the TI and goes beyond the traditional variogram-based (two-point) geostatistical simulation techniques. Compared to object-based modeling, the snesim is a pixel-based simulation method that provides more flexibility in conditioning facies simulation on hard and soft measurements. While the MPS presents several advantages over the two-point geostatistical simulation

methods and object-based modeling techniques, it requires a reliable qualitative description of the geologic continuity in the depositional environment (i.e., a TI). Moreover, the simulated facies inherit the important multipoint statistics that are conveyed by the TI (prior knowledge). This could be an advantage or liability depending on the quality of the TI.

A basic implementation of the snesim begins by using a search template to scan the patterns in the TI and convert them into a search tree. The search tree starts at a root node with the marginal probability of a facies type. It stores the transition probabilities by moving from the root node and considering possible pattern configurations (Strebelle 2002; Caers and Zhang, 2004). Once the TI is converted into a search tree, a random path is generated to sequentially visit and simulate facies types at un-sampled cells in the reservoir grid using the pattern probabilities stored in the search tree. The simulation proceeds by first obtaining the probability of a facies type at a simulation cell conditioned on the measured/simulated facies types in the neighboring cells from the search tree; it then uses the calculated probability to assign a facies type to the simulation node. Several options and facilities are included in the implementation of the snesim algorithm in order to account for various practical considerations and details during the simulation (Remy et al. 2009). In this dissertation, the snesim implementation in the Stanford Geological Modeling Software (SGeMS) was used (Remy et al. 2009).

1.3.1. Conditioning Facies Simulation on Soft Data

Conditioning MPS facies simulation on hard data is rather straightforward and is done by placing them in their corresponding simulation nodes and excluding them from the

random path of the simulation. However, the integration of soft data (e.g., three-dimensional seismic data) is more involved. The complication arises from the integration of multiple sources of data that can, in general, be dependent. Journel (2002) presents a method based on the permanence of updating ratios (i.e., the relative contribution of any single data event is assumed to be independent of any combination of other data events) for soft data integration that has been implemented in the snesim algorithm. Denoting a particular facies occurrence at a given location as event A and the facies probabilities based on the TI and soft data as $P(B)$ and $P(C)$, respectively, the conditional (or pre-posterior) probabilities $P(A|B)$ and $P(A|C)$ represent the facies probabilities conditioned on the TI (i.e., B) and on the soft data (i.e., C), respectively. The basic assumption in Journel (2002) is that the incremental contribution of data event C to the occurrence of A is the same before and after B is known. Journel (2002) shows that deriving the posterior probability $P(A|B,C)$ from the permanence of updating ratios principle is a more robust method for satisfying probability limit conditions than the traditional conditional independence hypothesis.

For the stated problem with two sources of data, the posterior probability is given by

$$\frac{x}{b} = \frac{c}{a} \tag{1.1}$$

where

$$x = \frac{1 - P(A|B,C)}{P(A|B,C)} \quad c = \frac{1 - P(A|C)}{P(A|C)} \tag{1.2}$$

$$b = \frac{1 - P(A|B)}{P(A|B)} \quad a = \frac{1 - P(A)}{P(A)}$$

The underlying assumption in Eq. 1.1 is equivalent to a form of independence (i.e., independence of relative incremental contributions) between data events B and C (Journel 2002). Data dependence can be introduced into this formulation in order to generalize the estimation of posterior probability using the τ-model (Journel 2002) as follows

$$P(A|B,C) = \frac{1}{1+x} \text{ with } \frac{x}{b} = \left(\frac{c}{a}\right)^{\tau(B,C)} \tag{1.3}$$

The value of $\tau \geq 0$ is interpreted as the weight given to data type C relative to data type B. For $\tau > 1$ ($\tau < 1$), data type C is given more (less) weight than data type B. Here, when conditioning facies simulation on production data, B represents the TI data and C to denote the probability map that is constructed using the production data.

Fig. 1.2 shows examples of conditional and unconditional snesim simulation results using a sample soft data (probability map). The TI used for this example is Fig. 1.2(a). The unconditional simulation uses a probability map, shown in Fig. 1.2(b), (left column); this map does not contain any information about the facies distribution, whereas for the conditional simulations shown in Fig 1.2(c), a probability map is used to constrain the spatial distribution of channel facies. It is clear that the information in the probability map strongly affects the spatial distribution of the simulated channel facies. In practice, the probability maps are usually obtained from the seismic data. In the proposed approaches, however, the probability maps were driven from the production data by using an inverse modeling approach.

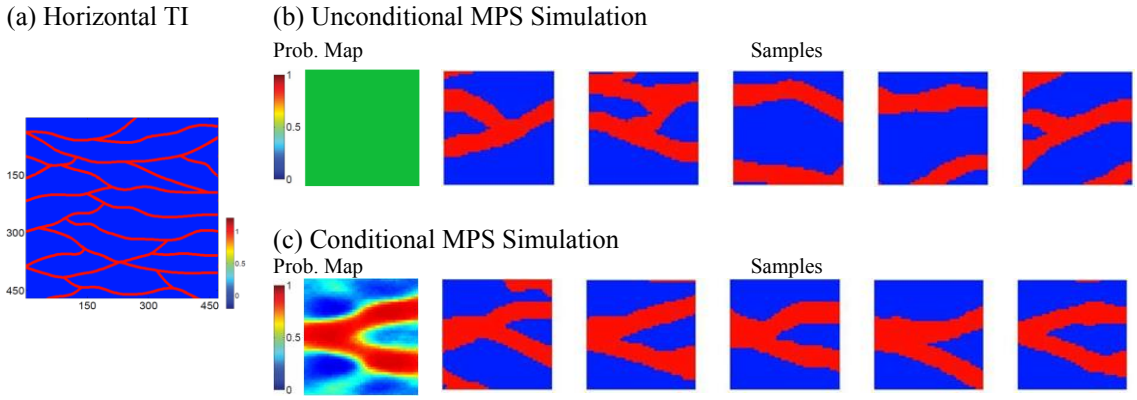


Fig. 1.2 – Conditional multipoint statistical facies simulation using soft data (probability maps). (a) A training image with mainly horizontal channels. (b) Uninformative probability map (left) and five unconditional samples of channels. (c) A sample probability map (left) and the resulting facies realizations conditioned on it. The simulation parameters (except for the probability maps) are identical in (b) and (c).

1.4. Data Assimilation Methods

In general, probabilistic interpretation of history matching problem leads to combining prior model parameter distribution, $p(m)$, with the new measurement data, d, to interfere likelihood model $p(m|d)$ in the framework of Bayesian approach. Because the relationship between the model parameters, m, and the observations is through nonlinear set of fluid flow through porous media equations, the posteriori probability distribution may not be easy to describe. In practice, except for specific simplified cases, complete description of the conditional density $p(m|d)$ is not feasible (Tarantola, 2005). Therefore, it is common to estimate some representative statistics, like mean and variance, of that conditional distribution using statistical inference approaches.

In this section, the production data integration approach, the ensemble Kalman filter, is presented.

1.4.1. Ensemble Kalman Filter (EnKF)

In the case of linear system, the Kalman filter (Kalman, 1960) is a recursive approach to estimate the internal state of the system from series of noisy observations. In contrast to batch estimators, the history of observations and/or estimates is not required here. The state-space model for the time-evolution of a linear dynamical system (ignoring the control term) can be expressed as:

$$\begin{aligned} x_{t+1} &= A_t x_t + w_t \\ d_t &= H_t x_t + v_t \end{aligned} \tag{1.4}$$

Here, the subscript t represents the current time step, x indicates the system state vector, A is the linear state transition matrix, d corresponds to the observation vector, H denotes the linear measurement model, and w and v stand for the modeling and measurement noises with covariances P and R, respectively. Eq. 1.4 could be used to obtain the mean and covariance of the state vectors in time according to the following assumptions:

$$\begin{gathered} E[x_0] = E[v_t] = E[w_t] = 0 \\ Cov[w_t, v_s] = Cov[x_0, w_t] = Cov[x_0, v_t] = 0 \\ Cov[x_0, x_0] = P_0 \qquad Cov[v_t, v_s] = R\delta_{ts} \qquad Cov[w_t, w_s] = Q\delta_{ts} \end{gathered} \tag{1.5}$$

The Kalman filter has two steps: the state vector prediction until the next observation step and the state vector updating with respect to the current measurements. The forecast step is:

$$x_{t|t-1} = A_{t-1}x_{t-1|t-1}$$
$$P_{t|t-1} = A_{t-1}P_{t-1|t-1}A_{t-1}^T + Q_{t-1} \tag{1.6}$$

Where, $x_{t|s}$ and $P_{t|s}$ refer to the estimate of the state vector and modeling noise covariance at time t conditioned on all observations up to and including time s. The update step includes the optimal linear combination of the forecast mean and measurements given by the following recursive equations.

$$\hat{x}_{t|t} = x_{t|t-1} + K_t\left(y_t - H_t\hat{x}_{t|t-1}\right)$$
$$\hat{P}_{t|t} = (I - K_tH_t)\hat{P}_{t|t-1} \tag{1.7}$$

Where K_t is the Kalman gain and it is defined as

$$K_t = \hat{P}_{t|t-1}H_t^T\left(H_t\hat{P}_{t|t-1}H_t^T + R_t\right)^{-1} \tag{1.8}$$

For large-scale nonlinear state-space problems (e.g. reservoir models), derivation of simple forecast equations for state mean and error covariance, like Eq. 1.6 for Kalman filter, is not trivial. Linearized versions of Kalman filter, such as the extended Kalman filter (EKF), have been used for small and medium scale problems with mild nonlinearities (Jazwinski, 1970). Because of the cost of linearization, the EKF methods are impractical for large scale and/or strong nonlinear problems.

In contrast to the linear system, the dynamical model for fluid flow through porous medium is described by a nonlinear system of differential equations and appropriate boundary conditions. Let assume $u_t \in R^{N_u}$ represents the discretized approximation of the solution at time t depends on poorly known parameters $m \in R^{N_m}$, need to be estimated. The dynamical model is:

$$\frac{du_t}{dt} = F(u, m) + \varepsilon_t^m \tag{1.9}$$

with initial state $u_{t_0} = u_0$. Here, ε_t^m and u_t represent the model error and all primary variables, e.g. pressure and saturations, in all grid blocks at time t, respectively. Also, the measurements at time $t_1, t_2, \dots, t_n$ with nonlinear relationship to the state variables

$$d_t = g_t(u_t, m) + \varepsilon_t^d \in R^{N_d} \tag{1.10}$$

where d_t and $g_t(\cdot)$ represent the measurement and forward model at time t, respectively and ε_t^d stands for the measurement errors.

In order to address the nonlinearity and dimensionality issues, Evensen (1994) provides a Monte Carlo approximation of Kalman filter as ensemble Kalman filter. Since then, the EnKF has been used widely in ensemble-based data assimilation problems (Houtekamer and Mitchell, 1998; Evensen, 1994; Reichle et al., 2002; Lorentzen et al., 2003; Nævdal et al., 2005; Wen and Chen, 2006; Skjervheim et al., 2007; Jafarpour and McLaughlin, 2009; Jafarpour and Khodabakhshi, 2011, Khodabakhshi and Jafarpour, 2013). In the EnKF, the probability density function (*pdf*) of the system state is represented using a collection of state vectors, called an ensemble. The covariance matrix of the Kalman filter is replaced by the sample covariance in the EnKF, and the Monte Carlo simulation is used to approximate the nonlinear time evolution of the states' forecast density.

In this dissertation, an augmented state vector containing reservoir parameters and simulated data is defined as,

$$x_t = \begin{Bmatrix} m \\ g_t \end{Bmatrix} \in R^{N_m + N_d} \tag{1.11}$$

Consequently, a linear relationship is established between the state vector and the data as:

$$d_t = Hx_t + \varepsilon_t^d \tag{1.12}$$

where, the matrix operator H extracts the simulated measurements from the full state vector. Let assume $x_{t_n:t_0}$ represent the sequence $(m, g_{t_0}, g_{t_1}, \dots, g_{t_n})$ and $d_{t_n:t_0}$ denotes the sequence $d_{t_1}, d_{t_2}, \dots, d_{t_n}$. Let also assume that the measurements at different time are independent and the forward model solution at a measurement time only depends on the solution at the previous measurement time, i.e. the forward model is first-order Markov process. Using the Bayesian statistics, the filter solution (Evensen, 2007) can be expressed as:

$$f(x_{t_n}|d_{t_n:t_1}) \propto f(d_{t_n}|x_{t_n}) f(x_{t_n}|d_{t_{n-1}:t_1}) \tag{1.13}$$

As a result, the model state and parameters with corresponding uncertainties are updated after each observation measurement. Finally, the ensemble-base version of Eq. 1.7 for the state vector update is written as (Evensen, 2003)

$$\hat{x}_{t|t}^j = x_{t|t-1}^j + P_{t|t-1}^e H^T \left(HP_{t|t-1}^e H^T + R_t\right)^{-1} \left(d_t^j - Hx_{t|t-1}^j\right) \tag{1.14}$$

where the superscript e indicates ensemble approximation. It is important to mention that the EnKF updates individual samples instead of updating first and second order moments, like original Kalman filter. Therefore, it is possible to use the updated replicates to approximate any desired (updated) statistics.

While EnKF has become increasingly popular for history matching continuous (and in particular, multi Gaussian) reservoir properties (Lorentzen et al. 2003; Nævdal et

al. 2005; Wen and Chen 2006; Skjervheim et al. 2007; Jafarpour and McLaughlin 2009a), it does not provide a mechanism for constraining the estimation of uncertain parameters on higher-order prior statistics. In addition, EnKF is a continuous estimator that is not suitable for the direct estimation of categorical variables (like facies). Nonetheless, the reservoir engineering literature has shown that the EnKF updates effectively integrate the flow data and provide important information about the spatial distribution of facies.

1.5. Research Objectives

The motive of this research is to extend the current state of state of art in detection of the large scale feature continuity during the process of history matching. In addition to theoretical developments, an effort is placed on the proposed approaches that are applicable to the real history matching problems.

The major parts of this research work are as follow. First in the Section 2, the probability conditioning method (PCM) for constraining multipoint statistical (MPS) facies simulation to nonlinear flow data is presented. The PCM implementation with the EnKF results in an improved performance of the filter updates, namely through the preservation of the facies correlation structure and the introduction of additional ensemble variability (spread) due to the resampling of facies from the TI after each update step. The important properties of the proposed PCM method are illustrated using several two-dimensional water flooding problems in reservoirs containing two and three facies types.

As an extension to Section 2, in Section 3, an adaptive strategy is proposed through a data feedback mechanism based on the predictive performance of the past realizations (i.e., sampling history) to improve the conditioning efficiency. The method is presented as a stochastic optimization algorithm and extended to the case where multiple training images are proposed as alternative plausible interpretation of the geologic scenarios for a given formation. Several numerical experiments from fluvial channel formations are used to demonstrate its applicability and performance in relation to non-adaptive conditioning techniques.

In Section 4, a novel Bayesian mixture model for adaptive and efficient sampling of conditional facies from multiple uncertain TIs is developed. The posterior distribution of facies is partitioned into individual conditional densities of the TIs and the corresponding mixture weights from the likelihood function for each TI are estimated. To implement the conditional sampling, the ensemble Kalman filter (EnKF)-based probability conditioning method is applied, whereby EnKF is used to invert the flow data and obtain a facies probability map (soft data) to guide conditional facies simulation from each TI. The suitability of the proposed Bayesian mixture-modeling approach is demonstrated using several numerical experiments in fluvial formations with uncertain orientation and structural connectivity.

In Section 5, a new workflow for integration of pressure data for estimating large-scale reservoir connectivity is proposed. According to the spatial resolution of the observations, the unstructured grid system is generated through Delaunay triangulation by using the location of the static pressure measurements as control points. The

important advantages of the proposed workflow for characterization of field-scale reservoir connectivity from pressure data include very fast connectivity estimation with a low-order model and effective parameterization to reduce the number of unknowns to a level commensurate with the available static pressure measurements. Successful application of the proposed approach to data from real fields illustrates its suitability and application to realistic reservoirs. The compatibility of the estimation results with the existing geological evidence demonstrates the performance of the proposed method.

In Section 6, the proposed workflow in Section 5 is applied for a real field data set to infer the large scale feature continuity using the available pressure data. Then, the finding in coarse scale history matching is downscaled to fine scale model and the probability conditioning method is applied to conclude the locations of channelized fluvial.

Section 7 summarizes this dissertation with the key finding in Section 2 to 6. Recommendations for future research are also presented.

2. PROBABILITY CONDITIONING METHOD (PCM)*

2.1. Summary

Probability conditioning method (PCM) is presented for constraining multipoint statistical (MPS) facies simulation to nonlinear flow data. MPS has recently been introduced for flexible grid-based simulation of spatial connectivity in formations containing discrete geologic objects (e.g., fluvial channels) that are not amenable to conventional two-point geostatistical modeling. Using the higher-order statistics in MPS, facies realizations are simulated from a conceptual geologic continuity model known as a training image (TI). As a result, the simulated realizations inherit the complex structural connectivity and multipoint spatial statistics conveyed by the TI. While conditioning multipoint simulation results on static hard (e.g., core) and soft (e.g., three-dimensional seismic) measurements is relatively straightforward, conditioning the simulated facies on nonlinear flow data is a nontrivial task. On the other hand, inversion methods that directly update post-simulation facies distributions have difficulty in reproducing the spatial connectivity (or higher-order statistics) implied by a TI. Using the PCM approach, the flow data is inverted to obtain a probabilistic spatial description of facies distribution (i.e., a probability map) and the resulting facies probability map is used to guide MPS facies simulation from a specified TI. Since the probability map

---

* Part of this section is reprinted with permission from "A Probability Conditioning Method (PCM) for Nonlinear Flow Data Integration into Multipoint Statistical Facies Simulation" by Jafarpour, B. and Khodabakhshi, M., 2011. Journal of Mathematical Geosciences. 43, 133-164. Copyright 2011 by Springer.

contains important information about the flow measurements, the simulated facies distributions are more likely to reproduce the observed flow data. While the proposed PCM approach can be used with different inversion algorithms, the ensemble Kalman filter (EnKF) is used to extract facies distribution probabilities from flow data. This choice was made because (i) the ensemble form of the EnKF is less sensitive to discontinuity and nonuniqueness (randomness) introduced in conditioning facies simulation on a probability map, and (ii) the EnKF has been established as an effective subsurface data assimilation approach.

The PCM implementation with the EnKF results in an improved performance of the filter updates, namely through the preservation of the facies correlation structure and the introduction of additional ensemble variability (spread) due to the resampling of facies from the TI after each update step. The important properties of the proposed PCM method are discussed and its effectiveness is illustrated using several two-dimensional water flooding problems in reservoirs containing two and three facies types. The conclusion is that PCM effectively combines the existing information in the flow data and the TI; it does so by using the former to infer probabilistic knowledge about inter-well and near-well spatial connectivity and the latter to ensure consistent facies structure and connectivity, where the flow data are inconclusive (e.g., away from measurement locations).

2.2. Dynamic Data Interpretation with Facies Probability Map

The uncertainty in the characterization and modeling of subsurface flow and transport properties presents a challenge in the modeling and design of a number of important environmental and energy applications, including groundwater flow and contaminant transport, energy recovery from geothermal and hydrocarbon reservoirs, and geologic storage of CO_2 in deep underground formations. An important component of predictive modeling is the description of the spatial variability in subsurface flow properties. Data sparsity and uncertainty in geologic continuity model result in significant uncertainty in the description of prior model input parameters. When formations containing predominantly discrete geologic objects are considered, the main characteristics of the fluid displacement patterns and production behavior are often controlled by the geometrical attributes and connectivity of these objects. Therefore, the history matching problem should be primarily focused on reconstructing the spatial distribution and continuity of these features.

A review of persevering large scale feature continuity was presented in Subsection 1.2. In this section, a novel approach is presented for conditioning MPS facies simulation on nonlinear dynamic flow data. The proposed approach is general and can be implemented using various nonlinear inversion techniques. The production data is used to infer the probability with which a particular facies type occurs at each grid block throughout the reservoir. Then, this spatial probability map is combined with the *snesim* algorithm to draw conditional facies realizations from a specified TI. The ensemble Kalman filter (EnKF) integrates the flow data and infer facies probability maps. While

EnKF has become increasingly popular for history matching continuous (and in particular, multi Gaussian) reservoir properties (Lorentzen et al. 2003; Nævdal et al. 2005; Wen and Chen 2006; Skjervheim et al. 2007; Jafarpour and McLaughlin 2009a), it does not provide a mechanism for constraining the estimation of uncertain parameters on higher-order prior statistics. In addition, EnKF is a continuous estimator that is not suitable for the direct estimation of categorical variables (e.g., facies). Nonetheless, it has been shown in the reservoir engineering literature that the EnKF updates effectively integrate the flow data and provide important information about the spatial distribution of facies. The EnKF estimates the mean of log-permeability from a prior ensemble and the flow data. To generate a facies probability map, the updated log-permeability mean is used with the snesim algorithm to simulate conditional facies realizations from a given TI. A review of MPS facies simulation and EnKF were presented in Subsections 1.3 and 1.4, respectively. This section is continued by a presentation and discussion of the results of the application of the method to several two-dimensional synthetic history matching examples with channel facies. Some of the important properties of the proposed PCM approach are discussed before presenting the conclusions.

2.3. PCM Implementation

The implementation begins by generating an initial ensemble of facies with an uninformative constant probability value of $\frac{1}{N}$ for N facies, which is equivalent to unconditional facies simulation from a prior TI. One probability map is generated for each facies type. After assigning permeability values to each facies, the facies ensemble

is used to predict the observed fluid flow quantities for the next time step. In the analysis step, the log-permeability realizations and the probability map are updated. The update equation requires up to the second-order statistics of the estimation parameters. The ensemble of log-permeability fields is used to derive the first- and second-order statistics required during the update step. The updated mean of the ensemble is then used to infer probabilistic information about facies distribution. The probability of each facies at a given grid block is assumed to be a function of the distance between the updated log-permeability value at that grid block and the log-permeability value of each facies. The analysis equation is only used to estimate the log-permeability mean from the production data. The implementation steps are given below.

2.3.1. Steps in Facies Conditioning

1. Generate an unconditional ensemble of facies using the *snesim* algorithm.

2. Start with an uninformative single probability map that assigns equal probability to each facies type (e.g., 0.5 for two facies).

Repeat steps 3–5 until all observations are processed.

3. Forecast step: assign permeability values to each facies in the model and run forward flow simulation, from initial time, to predict the states and measurements.

4. Update step:

   a. With the EnKF analysis equation for the ensemble mean, update the log-permeabilities using the new observations and find their mean map;

   b. Infer a new probability map from the updated log-permeability mean.

5. New facies generation: use the updated probability map with the *snesim* algorithm to generate a new ensemble of facies models.

The following remarks about the details of the implementation are necessary. First, after each update, a new probability map is constructed and used to generate an ensemble of conditional facies realizations. The state (pressure and saturation) solutions are obtained by rerunning the forward simulation from the initial time-step using the new facies models. Second, the update applied to the log-permeability fields is expected to be locally (at the vicinity of the wells) accurate and relatively inconclusive (or even potentially spurious) away from the wells. A probabilistic interpretation of such update results has an important implication for generating conditional facies maps: at the vicinity of wells where there is more confidence about facies distribution (very high/low probability of a particular facies), generation of facies will be more strongly affected by the updated probability maps. On the other hand, at increased distances from the wells, the updated probability map is likely to be inconclusive. Since the updated facies are simulated from the TI, the TI becomes more dominant, and the simulated facies inherit the structural continuity of the TI at locations in which the probability map does not provide conclusive information about the facies type. The exception to the above statement occurs when the probability map indicates (strongly) a different connectivity from that of the TI. In that case, the simulated facies may depart from the typical connectivity structure in the TI.

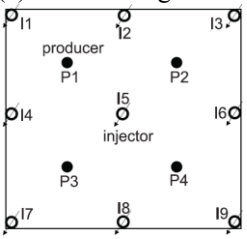
2.4 Applications

In order to motivate the proposed estimation approach, a two-dimensional water flooding example is used with a two-phase immiscible (oil/water) system and the EnKF is applied to infer the distribution of channel facies (resembling a fluvial environment) from production measurements. First, the EnKF directly updates log-permeability values in each grid block; this allows us to illustrate the inconsistency (with the prior) in applying the EnKF to discrete facies estimation and the degraded performance of the EnKF updates in time. The proposed PCM approach is applied to the same example. The discussion of the PCM method is continued with additional examples in order to evaluate its performance for the estimation of three facies and with TI uncertainty. This section is concluded by presenting more complex examples and by discussing the properties of the proposed approach.

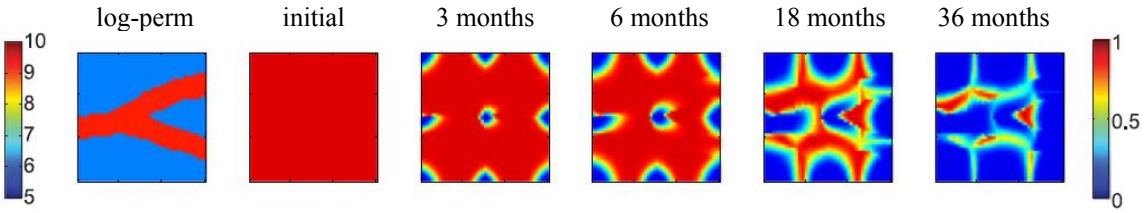
2.4.1. Experimental Setup

In the examples presented in the following section, four 5-spot pattern derives are combined to form a general 13-spot reservoir domain configuration consisting of nine water injection wells and four oil production wells (**Fig. 2.1(a)**). In each case, two-dimensional facies distributions are estimated from production measurements at well locations. Throughout, the only source of uncertainty is the facies distribution, so that all other model parameters are known. The reservoir is initially filled with oil at a constant uniform saturation of 0.90 (i.e., the connate water saturation is 0.10). The initial pressure in the reservoir was set at 6000 psi in all cells.

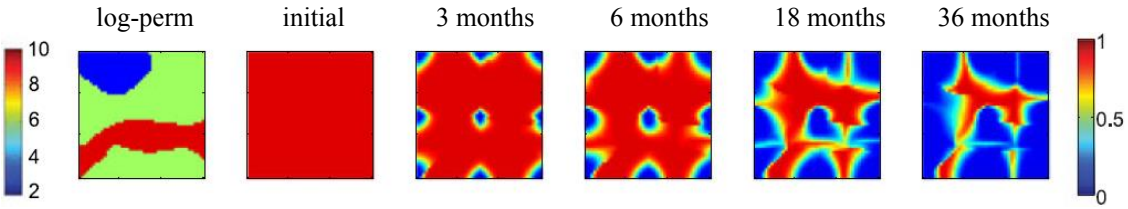
(a) Well configuration



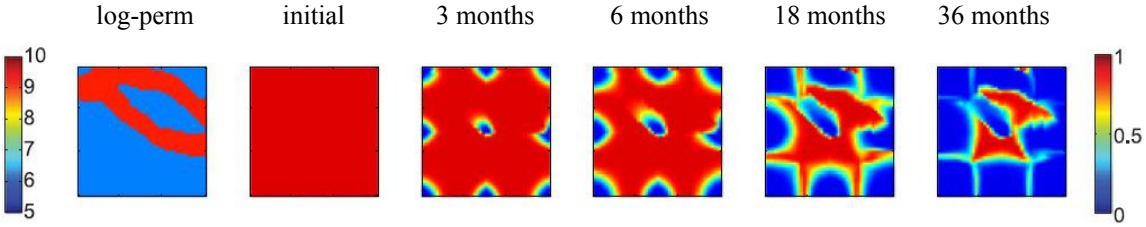
(b) Example 1: True facies and oil saturation



(c) Example 2: True facies and oil saturation



(d) Example 3: True facies and oil saturation



(e) Example 4: True facies and oil saturation

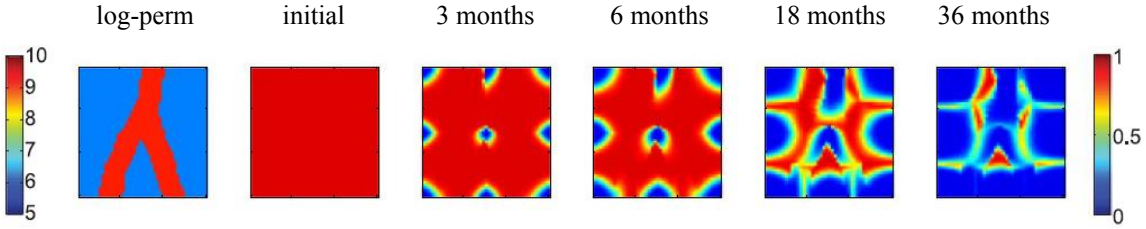


Fig. 2.1 – Field setup and the true saturation profiles. (a) A 13-spot reservoir configuration with nine injectors and four producers; (b) to (e) true log-permeability (left) and corresponding snapshots of saturation profiles in time (second to last column) for Examples 1, 2, 3, and 4, respectively. Simulation parameters are reported in Table 2.1.

Table 2.1 – General simulation and assimilation information

Simulation parameters			
Phases	Two-phase (oil/water)	Simulation time	12×3 months
Grid-system	45×45×1	Cell-dimensions	10 ×10×10 (ft^3)
Rock porosity	0.20 (constant)	Initial oil saturation	0.90 (uniform)
Initial pressure	6000 psi (uniform)	Injection volume	1PV
Number of injectors	9	Number of producers	4
Inj. wells constraints	Water flow rate	Prod. wells constraints	Pressure
Facies type	Fluvial formation	Geostatistical simulation	SNESIM
Assimilation parameters			
Obs. at injection wells	Pressure	Measurement noise	STD 20 psi
Obs. at production wells	Oil and water flow rate	Measurement noise	STD 20 sbpd
Observation intervals	3 months		

The injection ports are constrained to inject, with a uniform rate, a total of one pore volume (1 PV) during the thirty six months of simulation time. The production ports operate under pressure control, with a constant pressure of 5950 psi specified for each well. **Table 2.1** summarizes the general simulation and data integration parameters. Under these settings, the pressure at the injection ports, as well as the water and oil rates at the production ports, were measured, and used as the mean of observation distribution (Table 2.1), and assimilated. In each example, the measurements were obtained every three months by running a forward flow simulation with a specified (synthetic) true log-permeability field. Log-permeability values of 10,000 mD and 500 mD were assigned to the channel and background facies, respectively. Fig. 2.1(b) to 2.1(e) show the synthetic

true log-permeability models (left) and snapshots of the true saturation plots for Examples 2.1 to 2.4, respectively.

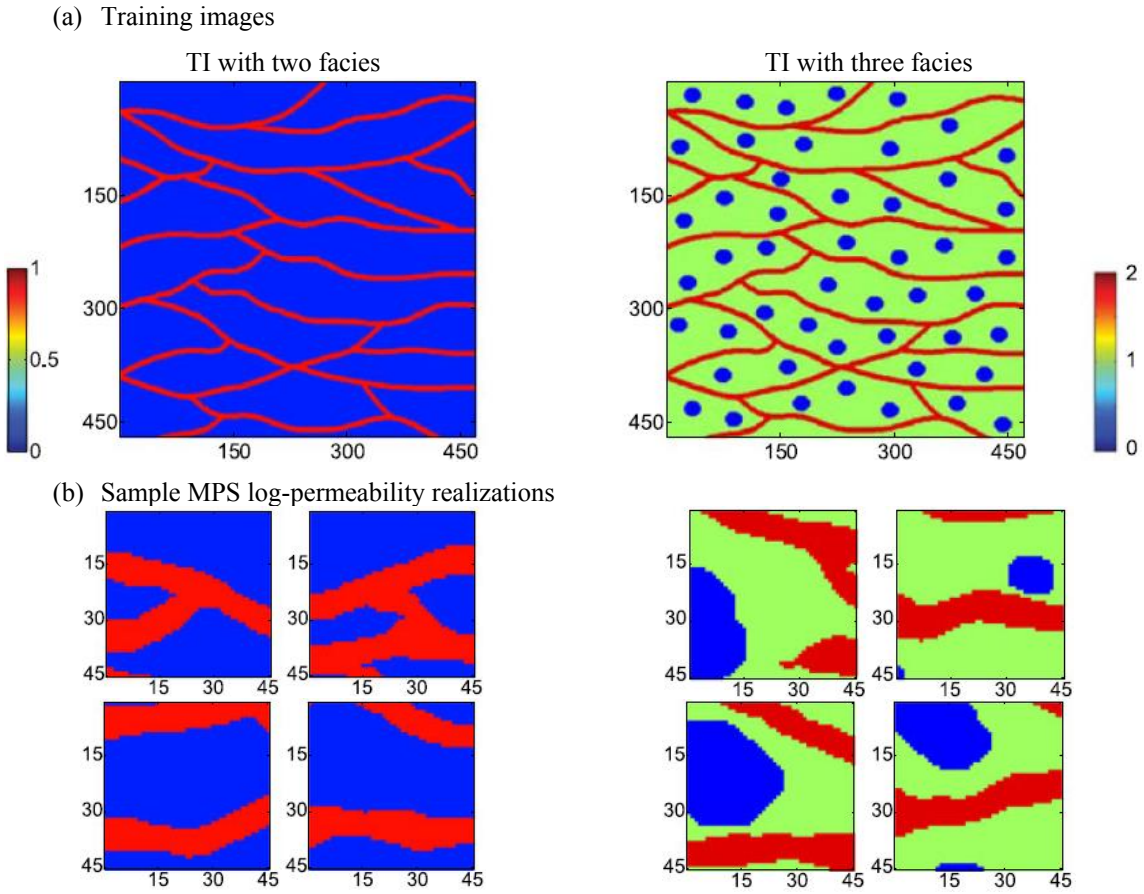


Fig. 2.2 – Training images and sample unconditional facies realizations for examples with two (left) and three (right) facies: (a) a 470 × 470 two-facies training image with horizontal channel orientations (left) and a similar training image with three facies (right), and (b) four unconditional realizations generated from the corresponding training images in (a) using the snesim algorithm.

2.4.2. Example 2.1: Estimation of Two Facies

2.4.2.1. Standard EnKF Solution

A history matching example is presented in which the EnKF estimates log-permeability values that represent the facies distribution. The synthetic true facies model in Example 2.1 is derived from the TI with two-facies (**Fig. 2.2(a)** (left)) and contains two horizontal channels that are connected in the middle of the domain. The TI is borrowed from (Mirowski et al. 2008). Four unconditional sample realizations from the same TI are shown in Fig. 2.2(b) (left). The simulated samples share the same statistics and general orientation (left to right) with the true model.

In all experiments, an ensemble size of N = 300 is used. The estimation results for Example 2.1 are shown in **Fig. 2.3**. Fig. 2.3(a) shows the true model, while Fig. 2.3(b) displays the initial sample log-permeabilities. Fig. 2.3(c) shows the history matching results obtained when the EnKF is used to update grid-block log-permeability values from flow data in thirteen observation locations. The log-permeability results after the first update show that the EnKF update equation treats the facies permeabilities as continuous random variables. Since the standard EnKF update equation does not provide any mechanism for the estimation of categorical variables, the estimated facies permeabilities do not preserve the categorical nature of the initial facies distributions. Examination of the results indicates that the marginal distribution of log-permeability at each grid block changes from a strictly bimodal (i.e., binary) to a Gaussian-looking distribution (not shown). It is clear from this example that the standard grid-based EnKF update may only be useful for local updating of facies distribution around the

observations. In general, the log-permeability results away from observation locations tend to be less informative about facies types. As the results in Fig. 2.3(c) suggest, applying the EnKF to update log-permeabilities at locations distant from the observed values can lead to spurious and inconsistent updates. In particular, when estimating categorical variables in which distinct features are to be reconstructed from limited observations, it is important to use an estimation approach that effectively combines local information conveyed by observations with global structural continuity (e.g., in the TI). The last row in Fig. 2.3(c) displays the changes in spatial distribution of the log-permeability ensemble variance over time. As expected, the variance decreases in time and increases as a function of the distance from the observation locations.

2.4.2.2. PCM Solution

In the PCM approach the log-permeability ensemble and its means are updated. Then, the log-permeability mean is converted to a single probability map using the mapping function shown in Fig. 2.4 (left). Here, a simple linear mapping assigns each updated mean log-permeability value to a proportional facies probability value within the specified bounds. Fig. 2.3(d) shows the facies estimation results obtained after the first and last updates in Example 2.1, using the proposed PCM approach. The first row contains the updated probability maps and shows that the probability map reveals a stronger signature of the true facies distribution as observations are integrated in time. The final facies distribution after the integration of all measurements appears to capture the overall orientation and connectivity of the channel features in this example.

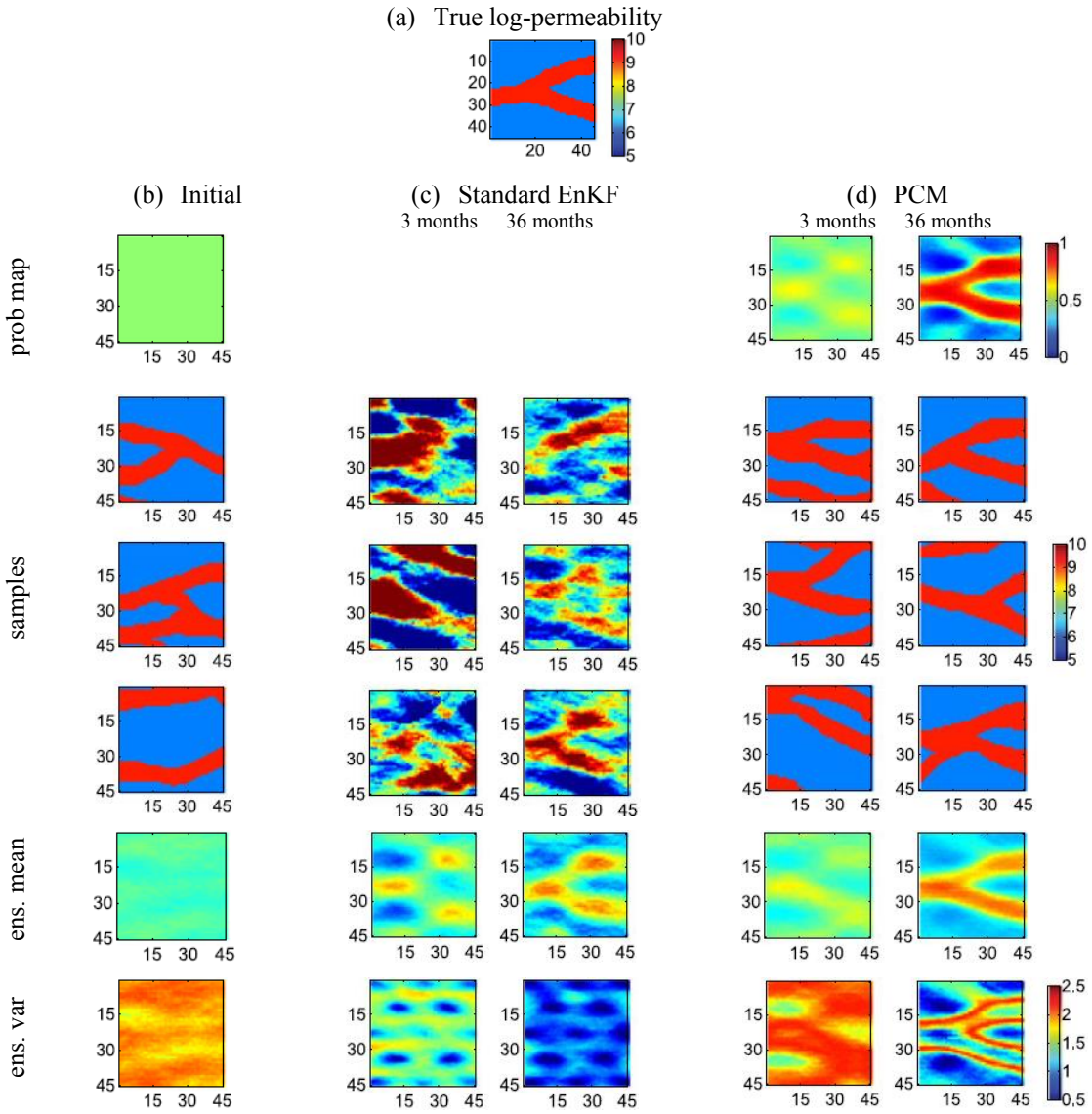


Fig. 2.3 – Facies estimation results for standard EnKF and the PCM approach in Example 2.1. (a) True log-permeability field; (b) initial probability map (top), three sample log-permeabilities (second to fourth rows), and the full ensemble log-permeability mean and variance (fifth and sixth rows, respectively); (c) three sample log-permeability replicates (corresponding to initial permeabilities realizations in (b)) after three months (left) and thirty six months (right) updated with standard EnKF (first to third rows); the updated log-permeability ensemble mean and variance are shown in the fourth and fifth rows; (d) the probability maps (first row) and the same three sample log-permeability replicates after 3 months (left) and 36 months (right) updated using PCM (second to fourth rows); the updated log-permeability ensemble mean and variance are shown in fifth and sixth rows.

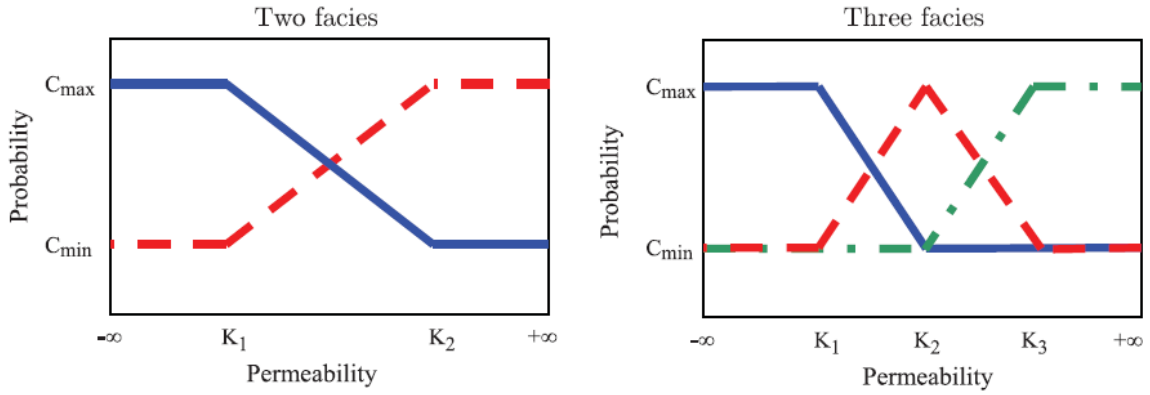


Fig. 2.4 – Functions used to convert mean log-permeability fields to probability maps for the case with two (left) and three (right) facies.

The results in Fig. 2.3(d) are clearly distinct from those in Fig. 2.3(c) in important ways. Firstly, throughout the assimilation the updated facies always follow a categorical distribution after conditioning on production data, which is consistent with the prior information. This consistency has important implications for the improvement of the sample statistics needed for future updates (discussed below). Secondly, the resampling of the facies ensemble from the TI after each update step introduces additional variability and spread to the ensemble. This helps to reduce the effect of variance underestimation or ensemble collapse. It is important to note that the level of variability in the generated ensemble can be controlled by either the limits specified for the probability map and/or by specifying the τ value in Eq. 1.3. Here, the bounds $[0.10, 0.90]$ on the probability map only allow for a 90% confidence on the updated results. Note that even with a relatively confident probability map, a reasonable amount

of variability and spread is introduced into the ensemble. Changing the bound to $[\epsilon, 1-\epsilon]$, where $\epsilon < 0.1$, can result in reduced ensemble variability.

An interesting observation can be made from the results in Fig. 2.3(d): while the updated probability maps provide little connectivity at the early time steps, they tend to capture the facies connectivity at later assimilation time steps. Unlike the results in Fig. 2.3(c), the connectivity in the probability map is preserved much better when facies are resampled from the TI. This result is mainly attributed to the more accurate and consistent conditional facies maps that are generated from the TI at each update step. In the standard EnKF for estimation of log-permeability, the original continuity structures break apart after the first update step and the resulting sample statistics quality decays. Hence, the sample covariances used for future updates are less accurate, which adversely affects the estimated log-permeabilities at later times. In contrast, when probability maps are used, the ensemble of facies maps are regenerated from the same TI using additional production information. Hence, the resulting covariances are more accurate than the initial covariance. As a result, the updated probability maps and resampled facies tend to improve with time. A key advantage of the proposed approach is the consistency of the forecast covariance in the representation of the channel-like structures. While second-order statistics are not adequate for the representation of categorical facies, regenerating facies from the TI after each update step yields a more accurate approximation of the facies covariance matrix and hence improved updates over time.

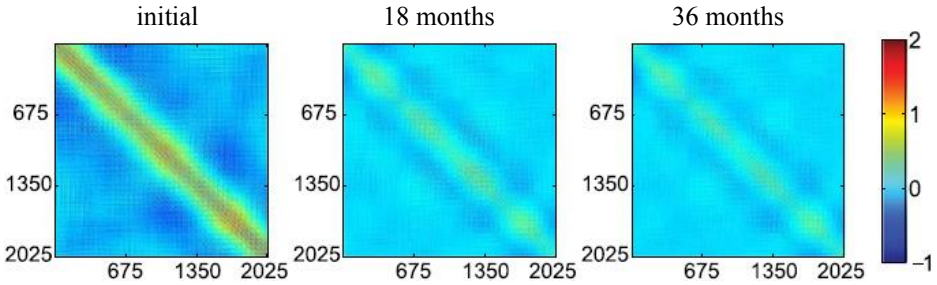
To highlight this important difference, **Fig. 2.5** illustrates the initial covariance matrix (left) and its evolution after six and twelve updates at eighteen (middle) and thirty

six (right) months, respectively. These covariances are generated through row-wise vectorization (along channel directions) of the facies maps. The top 45×45 diagonal block sub-matrix therefore contains the covariance values between forty five gridblocks in the first row, while the next 45×45 diagonal block matrix represents the covariance between the grid-blocks in the second row, and so on. As such, the first off-diagonal 45×45 block matrix (that is, with indices ($i = 1:45, j = 46:90$)) contains the covariance between the grid blocks in the first and second rows of the domain. Therefore, the number of off-diagonal red bands in Fig. 2.5 indicates the covariance across the width of the channel. The covariance update results are shown for both grid-block log-permeability estimation (Fig. 2.5(a)) and facies estimation with probability map using the PCM approach (Fig. 2.5(b)). The comparison of the covariance matrices for the two cases makes it clear that the original correlation structure that captures the large-scale continuity is lost during the grid-based log-permeability update steps.

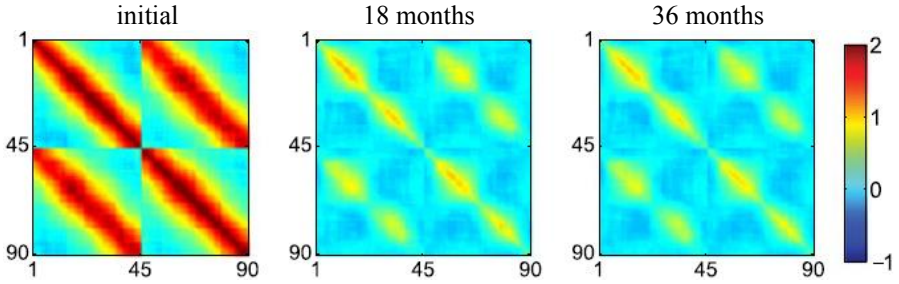
A closer examination of the enlarged 90×90 top-left diagonal block covariances (corresponding to the second rows in Figs. 2.5(a) and 2.5(b)) clearly indicates that the covariance matrices generated from the conditional facies after update steps preserve the original continuity structure better.

Another important difference between the two results lies in the estimated variance after the updates. A comparison between the last rows of Figs. 2.3(c) and 2.3(d) shows that a large portion of the variance in the conditional facies samples is concentrated around the edge of the channels and away from the observations. This is mainly because of the variability introduced by the TI-based sampling.

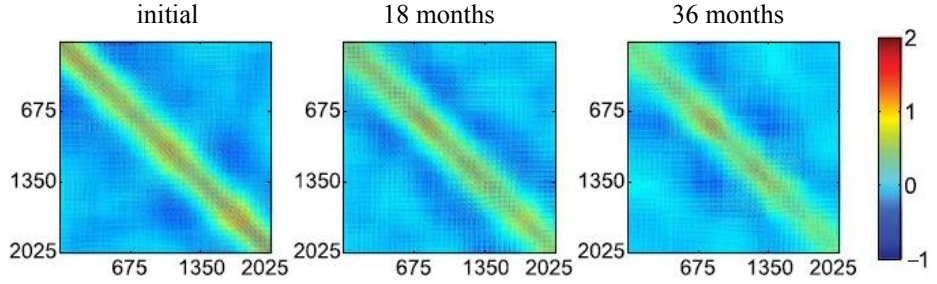
(a1) Initial and updated log-permeability covariances (grid based update Example 2.1)



(a2) Zoomed in version of (a1) with the first 90 × 90 elements



(b1) Initial and updated log-permeability covariances (prob map update Example 2.1)



(b2) Zoomed in version of (b1) with the first 90 × 90 elements

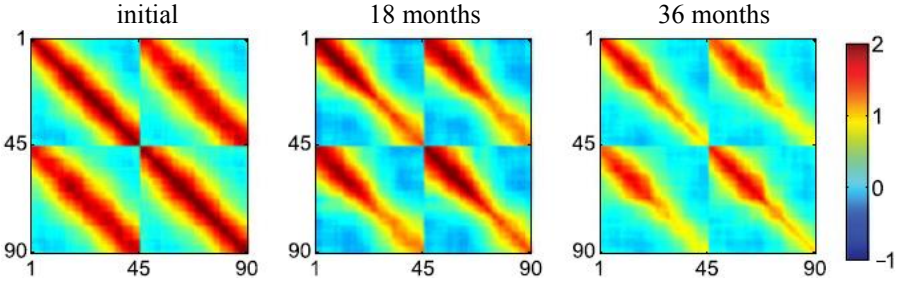


Fig. 2.5 – The initial and updated covariances in Example 2.1. (a1) and (a2) show the initial covariance (left) and the updated covariance after eighteen (middle) and thirty six (right) months when grid-block log-permeability values are updated; (a2) shows a zoomed-in version of the top-left 90×90 sub matrices. (b1) and (b2) show similar plots when the probability maps are updated and used in the snesim algorithm to generate conditional facies samples.

On the other hand, when log-permeability values are directly updated (Fig. 2.3(c)), the variance increases with the distance from the observation locations. Moreover, less variability is associated with channel edges, which had disappeared after the updates.

Fig. 2.6 provides the summary of the production plots for Example 2.1. The top row figures show the predicted oil rate, while the bottom row displays the water rate plots for the unconditional (dashed black lines), grid-based (solid blue lines), and probability map-based (solid green lines) facies estimation. While the true production plots are not covered by the initial ensemble of forecasts, the update results with PCM move the forecasts closer to the true reservoir responses. The estimation of the full ensemble response shows that, overall, the production forecast plots resulting from estimation with the standard EnKF tend to underestimate the uncertainty (variability) in the facies distribution, whereas more variance is observed in the production forecasts corresponding to facies estimated using the PCM implementation.

Since the standard EnKF update equation is not constrained to honor the flow equations, the updated states and parameters do not satisfy the mass and momentum conservation principles represented by multiphase flow equations. Two general positions with regard to this issue have been taken in the literature. The first view suggests using the updated pressure and saturation states to initialize the forecast for the next time step. The underlying assumption in this approach is that the flow equations are imperfect because of the simplifying assumptions and errors introduced in deriving them. Hence,

violation of the conservation principles should be tolerated. The alternative view emphasizes the conservation principles by restarting the simulations from the initial time step and by using the updated parameter ensemble to derive a consistent ensemble of current states for initializing future forecasts. This step, while ensuring consistency, comes at the cost of additional simulation time. In our PCM implementation, the resampling step from the TI generates new facies realizations that are not consistent with the EnKF updated saturations and pressures. A restart step derives the pressure and saturation states from the updated facies distributions.

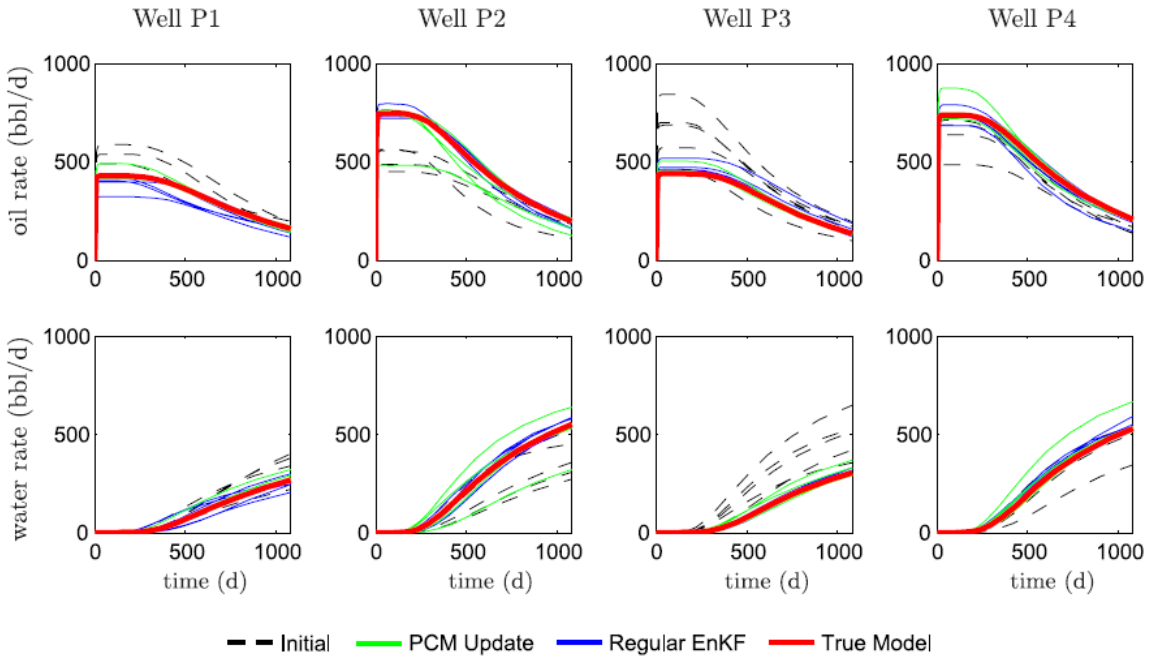


Fig. 2.6 – Production forecasts for Example 2.1. Oil production rate (top) and water production rate (bottom) forecasts corresponding to the five sample log-permeability fields (in Fig. 2.3) are shown for the initial (dashed lines), updated with standard EnKF (solid blue lines), updated using PCM (solid green lines) and the true (solid red line) log-permeability fields.

The overly good production forecasts (with lower ensemble spread) in the case of grid-based log-permeability estimation reveal an important fact about the non-uniqueness of the inverse modeling solutions. Evidently, the estimated permeabilities are far from the true facies, yet the production forecasts reproduce the true production behavior very well. The implication of this observation is quite important in that it cautions against putting too much emphasis on just fitting the production data as a metric for estimation performance. This can be better appreciated by observing that the saturation profiles may show different sweeping trends from those of the true saturations. As a result, the production curves can be matched by a wide range of log-permeability estimates that may not share the correct log-permeability structure and saturation distribution. Although accurate production forecast plays an important role in characterizing and controlling the input–output behavior of the reservoir, it does not provide information about the sweep efficiency of the water flood and the spatial distribution of fluids in the reservoir. The latter is critical for the planning of future reservoir development, activities such as drilling new infill wells and placing additional wells to produce any by-passed oil.

An important element contributing to the relatively larger variance in the production plots of the proposed PCM approach is the randomness introduced during the snesim simulations. Samples drawn from the TI are usually more diverse than the updated samples in the regular EnKF implementation; this variability is particularly higher in areas away from observations where facies probabilities are inconclusive. As a result, the production forecasts from the PCM approach have a wider spread. The

variability introduced in the ensemble can improve ensemble collapse issues by having an effect similar to covariance inflation, which is a common practical measure for dealing with the EnKF ensemble collapse. On the other hand, the sampling variability can introduce a small fraction (about 10% in our case) of outlier samples that exhibit completely incorrect production behavior. These samples can be removed from the ensemble by using an importance sampling technique (Robert and Casella, 2005).

2.4.3. Example 2.2: Estimation of Three Facies

To investigate the applicability of the proposed PCM approach to the estimation of multiple facies distribution, an example with three facies is used. A low permeability facies is added to the first TI to generate a new TI with three facies, namely high permeability (sandstone channels), medium permeability (background), and low permeability (salt dome). The new TI is illustrated in Fig. 2.2(a) (right). The true facies model is obtained from the same TI. The true model (Fig. 2.1(c), left) and saturation profiles at the initial time step and after three, six, eighteenth, and thirty six months are presented in Fig. 2.1(c) (second to sixth columns, respectively). Four initial facies samples (out of the three hundred) are depicted in Fig. 2.2(b) (right). The permeability value of the new facies (salt dome) is 20mD. All other simulation and data integration parameters are the same as in the previous example. To convert the updated log-permeability mean into the appropriate facies probability map, the simple functional relation illustrated in Fig. 2.4(b) is used. The limits for each probability map are set to [0.05, 0.9]. For the sake of comparison, the grid-block log-permeability values are updaed. The results obtained from both methods are shown in **Fig. 2.7**.

Fig. 2.7(c) illustrates the grid-block log-permeability estimation results obtained using the standard EnKF. Two initial facies replicates, as well as the corresponding updates after six and thirty six months (second and last updates) are displayed in the second and third rows. The fourth and fifth rows depict the mean and variance of the ensemble after each update. Fig. 2.7(d) shows the facies estimation results for the second example using the PCM approach. The second and third rows show the time evolution of the probability map for channel and salt dome facies, respectively, after the assimilation of the production data. The initial probability map is completely uninformative since it implies that the initial facies maps were generated using only the TI information. However, as flow data are integrated through the EnKF update steps, the probability maps begin to reveal signatures of the true facies distributions. The final facies distributions appear to capture the overall location, orientation, and connectivity of the true channel as well as the location and size of the salt dome. The fourth and fifth rows show two out of three hundred replicates drawn from the TI after conditioning on the updated probability maps. The last two rows display the mean and variance of the log-permeability ensemble for the corresponding update step.

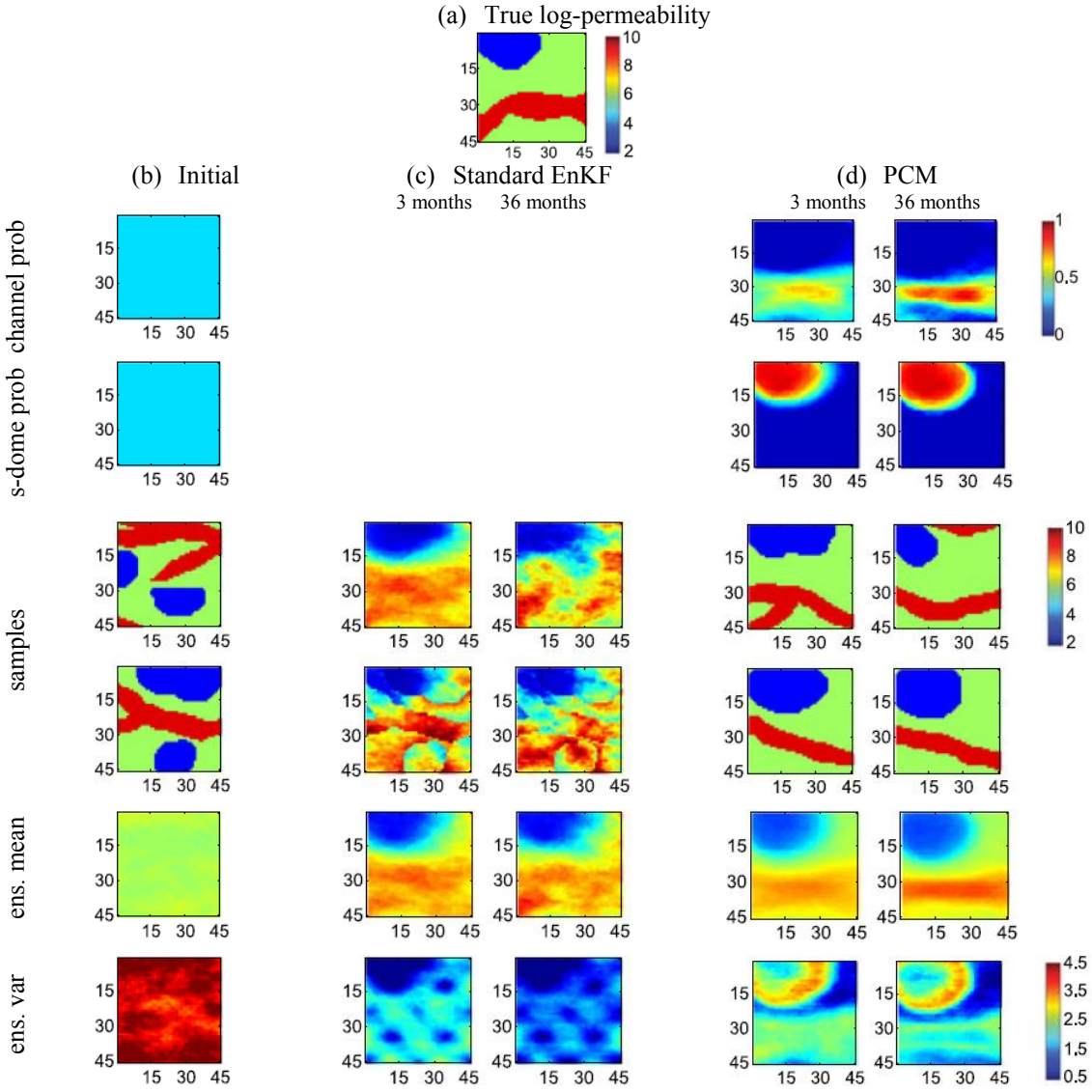


Fig. 2.7 – Estimation results for Example 2.2 with three facies: (a) the true log-permeability model; (b) initial probability maps for channel (second row) and salt dome (third row) facies, two sample log-permeabilities (fourth and fifth rows) and the full ensemble log-permeability mean and variance (sixth and seventh rows, respectively); (c) estimated log-permeability results with the standard EnKF and after three (left) and thirty six (right) months; two sample log-permeability replicates (corresponding to initial permeabilities realizations in (b)) after three (left) and thirty six (right) months (second and third rows) as well as the corresponding updated log-permeability ensemble mean and variance (fourth and fifth rows); and (d) updated probability maps for channel (second row) and salt dome (third row) facies after three (left) and thirty six (right) months using PCM; the updated sample log-permeabilities (fourth and fifth rows) and the ensemble log-permeability mean and variance (sixth and seventh rows, respectively).

The results obtained from grid-block log-permeability updating Fig. 2.7(c) are markedly different from those obtained using the PCM approach (Fig. 2.7(d)). When log-permeability values are estimated, the connectivity and structure of the facies degrade after the first update. On the other hand, since the log-permeability realizations in the PCM approach are drawn from the TI, the updated facies are consistent with the prior information. Fig 2.7(d) shows that the main uncertainty is distributed along the edges of the salt dome and channel facies. This uncertainty is attributed to the resampling of the facies from the TI after each update step. The shape of these features acts as a contributing factor to the higher uncertainty in the estimation of the location and size of the salt dome. The channel facies connect the left and right ends of the reservoir domain and pass through more wells. As a result, they are easier to observe than the salt dome facies. Overall, the superior performance of the PCM approach in reconstructing the spatial facies distribution from the flow data is also evident in this example. The production forecasts for this example (not shown) were consistent with the discussion presented for Example 2.1. Next, the performance of the PCM method under uncertainty in the TI is considered.

2.4.4. Uncertainty in Training Image

2.4.4.1. Example 2.3: Consistent TI

In previous examples, it was assumed that the TI provides a representative prior model for the synthetic true facies. The validity of this assumption for realistic facies estimation problems is open to question. Uncertainty in the TI is an important issue in the field application of multipoint geostatistical simulation. With the following two examples, the

effect of training-image uncertainty on the performance of the PCM algorithm is briefly discussed. A slightly more complex true facies model is used to motivate the discussion on the TI uncertainty. The true facies model in this example is derived from the same TI; however, the underlying feature in the simulated true channel is not commonly found in the TI. **Fig. 2.8(b)** shows the estimation results for this case. Since the TI is not a good prior model for representing the true facies distribution, the estimation problem becomes more challenging. Fig. 2.8(b) clearly shows that, although the approximate location of the channel features is detected, the estimated facies do not capture the exact shape of the true channels. The updated probability maps in this case are not as good as those in Example 2.1, mainly because the simulated facies (and the covariances generated from them) do not include the shape and geometric attributes of the true model. In spite of this fact, the probability map provides useful information about the location and approximate shape of the channel features. The inaccuracy in the training image seems to produce simulated facies that do not resemble the true model, and thereby contributes to the degradation of future updates.

Note that, even though the TI fails to simulate facies that share the true features, the updated probability map tends to improve with time. In the standard EnKF, when process noise is not included, the lack of a representative prior model can lead to filter divergence. In these cases, the ensemble spread is rapidly reduced while the solution diverges from the true model (Brouwer and Jansen 2004; Jafarpour and McLaughlin 2009a). Ensemble collapse does not seem to be a major issue in our implementation because ensemble variability is reintroduced into the forecast step through simulation of

new facies at each snesim sampling step. The results obtained in this example, however, point to the importance of a reliable prior TI for achieving successful performance in the related history matching problems. Even in the case where the TI and true model are statistically consistent, it appears that the lack of sufficient representative features in the TI can lead to degraded estimation performance. In some cases, the structural difference between the updated properties and the dominant features of the TI may provide useful guidelines for modifying the TI and repeating the inversion procedure.

2.4.4.2. Example 2.4: Inconsistent TI

Our final example in this section illustrates an extreme case of an inconsistent TI. The true model in this example is the clockwise rotated version of the true facies from Example 2.1. **Fig. 2.9(a)** shows the true facies model and the corresponding saturation profiles. The TI used to draw the ensemble of facies is the same as in Example 2.1, which is completely inconsistent with the true facies model.

Not only are the grid log-permeability estimation results shown in Fig. 2.9(b) disconnected, they also fail to correct the inconsistent direction of the initial channels throughout the updates. Previous studies on the EnKF performance under incorrect (biased) initial geologic continuity (Jafarpour and McLaughlin 2009a; Jafarpour and Tarrahi 2010) have shown that EnKF can have difficulty in fixing an incorrect (biased) structural property in the initial ensemble.

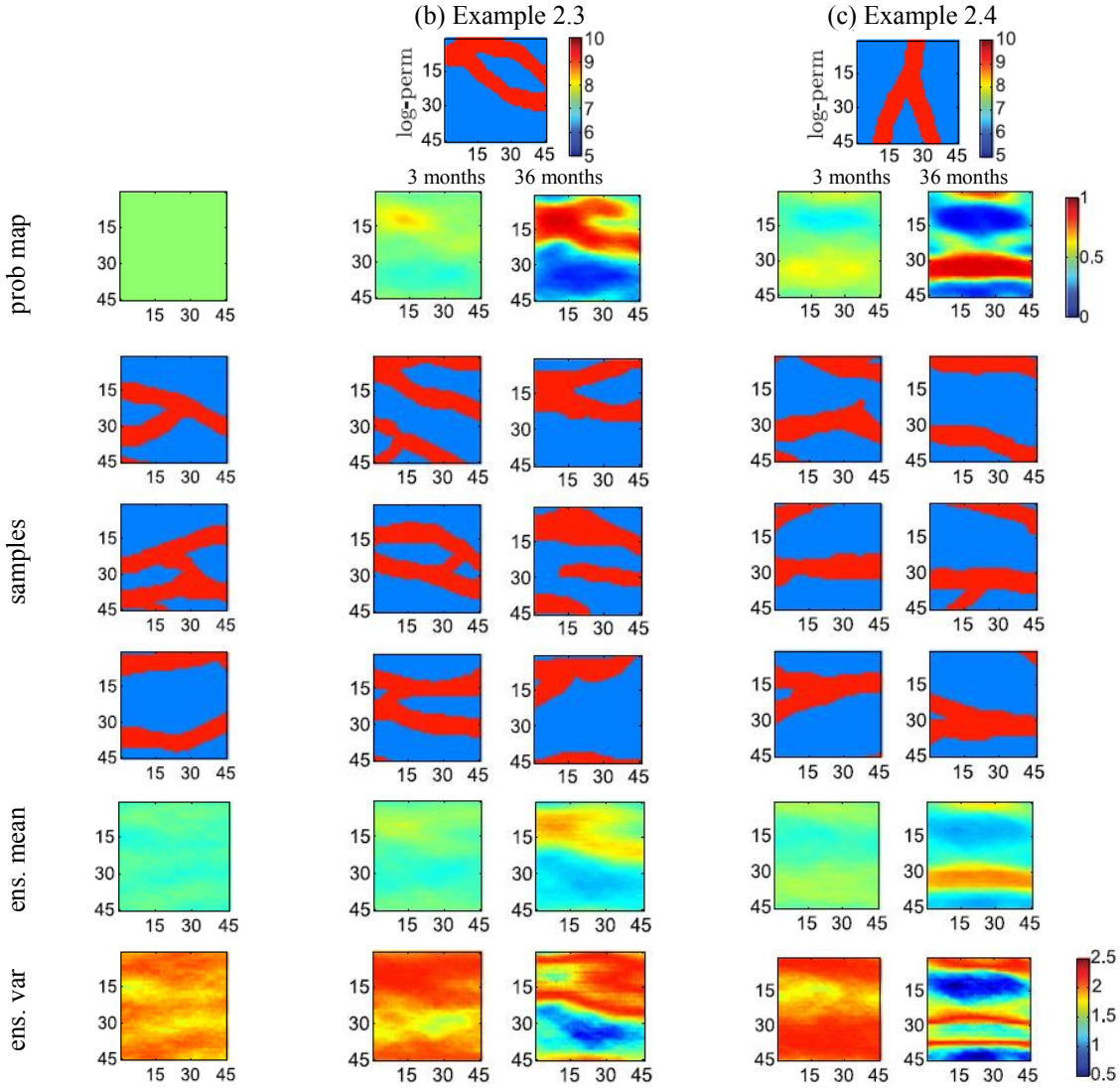


Fig. 2.8 – Facies estimation results for Examples 2.3 and 2.4. (a) Initial probability map (top), three sample log-permeability samples (second to fourth rows), and the full ensemble log-permeability mean and variance (fifth and sixth rows, respectively). (b) Estimation results for Example 2.3 using the PCM approach; the top panel shows the true log-permeability, the first row of bottom panel displays the updated probability maps after three (left) and thirty six (right) months, the second to fourth rows of the bottom panel show three sample log-permeability replicates; the updated log-permeability ensemble mean and variance are shown in fifth and sixth rows. (c) Estimation results for Example 2.4 using the PCM approach; the top panel shows the true log-permeability, the first row of bottom panel displays the updated probability maps after three (left) and thirty six (right) months, while the second to fourth rows show three sample log-permeability replicates; the fifth and sixth rows show the updated log-permeability ensemble mean and variance.

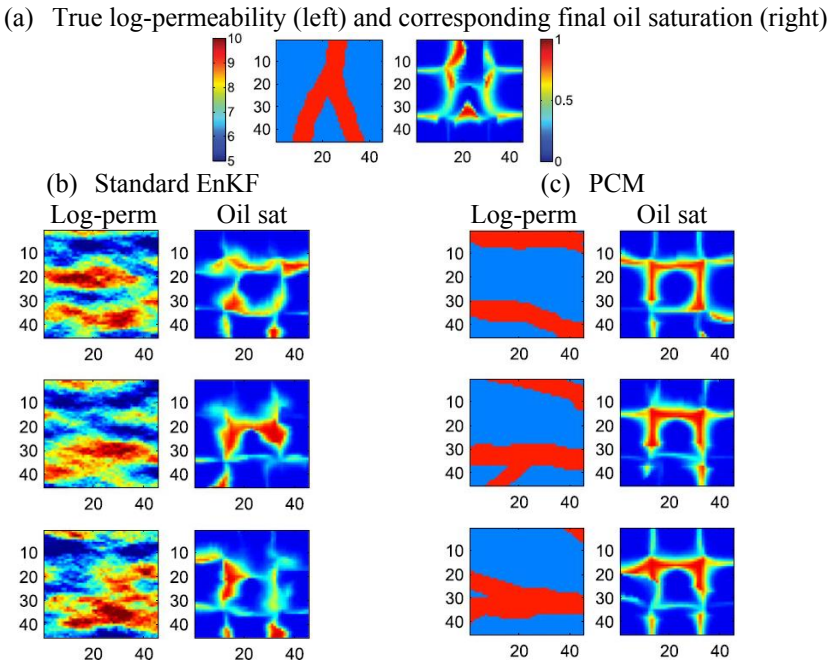


Fig. 2.9 – Final log-permeability and oil saturation results for Example 2.4: (a) true log-permeability (left) and the corresponding final true saturation profile; (b) sample final log-permeability estimates for the standard EnKF (left) and their corresponding final oil saturation predictions (right); and (c) sample final log-permeability estimates from the PCM approach (left) and their corresponding saturation forecasts (right).

The resulting production forecasts for the sample permeabilities in Fig. 2.9(b) are shown in **Fig. 2.10** (blue lines). Interestingly, despite the poor log-permeability estimation results, the production forecasts are very close to the data from the true model. While this may appear surprising, it can be explained by the ill-posed nature of the problem and by the existence of many non-unique solutions. It is relatively easy to history match production data if the only objective is to match the production history. However, the main goal of the history matching practice is to generate accurate

predictive models that honor all observations and prior geologic information. Fig. 2.9(b) shows the final saturation profiles corresponding to each sample log-permeability field. A comparison of the true saturation profile with those from the sample permeabilities reveals that the by-passed oil regions are in fact quite different. Unfortunately, the flow data from scattered wells alone do not seem to provide any indication of inconsistency in the prior ensemble. This is an issue that can be improved by incorporating additional data types with more extensive coverage (e.g., four-dimensional seismic).

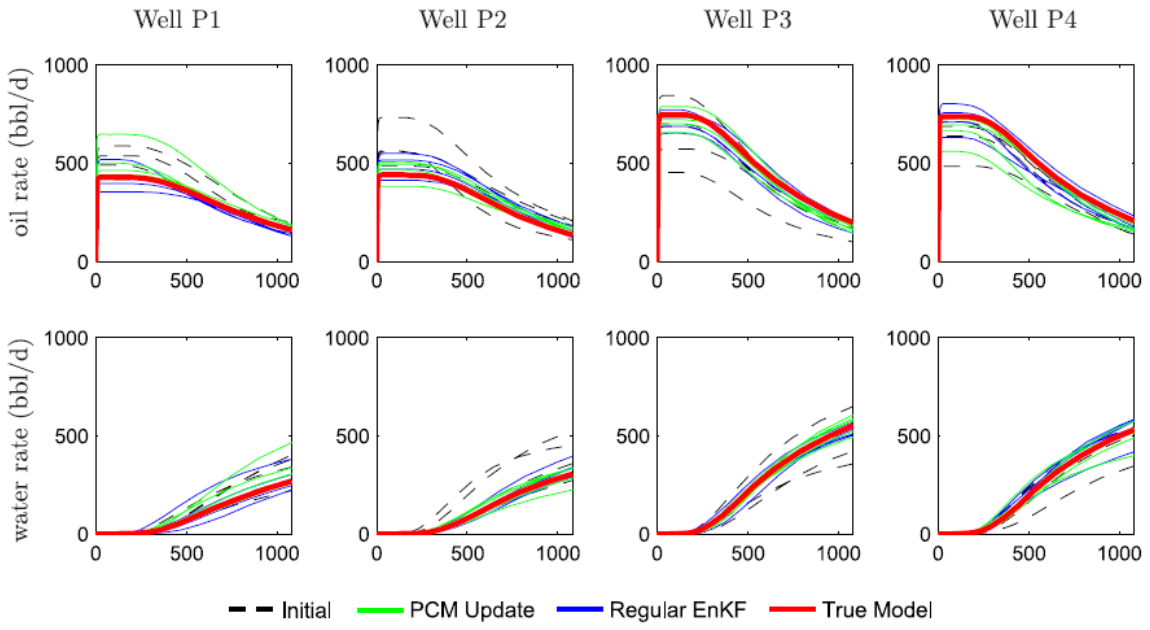


Fig. 2.10 – Production forecasts for Example 2.4. Oil production rate (top) and water production rate (bottom) predictions corresponding to the five sample log-permeability fields are shown for the initial (dashed lines), updated with standard EnKF (solid blue lines), updated using PCM (solid green lines), and the true (solid red line) log-permeability fields.

The facies estimation results with the PCM approach are shown in Fig. 2.8(c). It can be clearly seen that the initial horizontal channel features tend to persist during the EnKF updates. The final facies ensemble consists of samples that are significantly different from the true facies. The effect of an incorrectly specified prior ensemble is evidently too strong to be corrected by the limited available observations in this example. The well flow response and saturation prediction results are also shown in Figs. 2.9(c) and 2.10. These figures clearly show that while the flow predictions provide reasonable matches to the observed data, the saturation results are not similar to the true saturation profile. The proposed PCM approach cannot correct the bias in an incorrect prior, an outcome that highlights the significance of a reliable TI in conditioning facies simulation.

This example clearly shows that the uncertainty in the TI is an important and rather detrimental element in TI-based history matching problems and that it is one that deserves further investigation. The proposed approach can be implemented namely to address TI uncertainty by using several TIs that reflect alternative plausible geologic scenarios that might occur in the field. In that case, an effective sampling strategy should be incorporated to adaptively include more samples from the consistent TI. This is an issue that will be addressed in the next section.

2.5 Conclusion

A new probability conditioning method was presented for constraining TI-based facies simulation to reproduce dynamic flow measurements. While the EnKF update equation

was employed for production data integration, the PCM approach is quite general and can be implemented by means of alternative inverse modeling techniques. Since discrete geologic facies contain features that are more amenable to description with categorical random variables, a standard grid block property estimation with the EnKF can lead to inconsistent facies connectivity and spurious updates, especially at distant locations from observations. Although EnKF has been shown to be an effective history matching method, the continuous, second-order, and unconstrained nature of its update equation is poorly suited to the estimation of discrete geologic objects (e.g., fluvial channels) which can only be accurately described if higher-order statistical moments are included. The proposed probability conditioning method uses the EnKF to update the map of log-permeability mean from the production data (and a prior ensemble of models). This is then used to infer a probabilistic description of facies distribution. The resulting probability map is incorporated into TI-based multipoint geostatistical simulation algorithms in order to generate conditional facies realizations. Facies resampling (from the TI) after each EnKF update step offers several advantages over the standard implementation of EnKF. Some of these advantages include the following: (i) consistency of the updated facies with the prior TI (geologic model) and preservation of higher-order statistics; (ii) probabilistic interpretation of the production data to reduce the effect of spurious EnKF updates; and (iii) improvement of the ensemble variance underestimation (and collapse) issues through the introduction of additional ensemble variability, which is done by resampling new facies from the TI after each update step.

The implementation of the proposed framework was presented and discussed, along with synthetic two-dimensional water flooding examples with different levels of TI representativeness. It was shown that the proposed framework could effectively incorporate the production data into conditional facies simulation using the snesim algorithm. The production forecasts that were obtained when the proposed approach was used provided wider ensemble spread than those obtained from the estimation of grid-based log-permeability maps, mainly because of the variability and randomness introduced by resampling new facies from the TI. Occasionally, the randomness in the snesim facies simulation yielded a small number of obvious outliers that were likely to be rejected in an importance sampling study.

An example with three facies types was used to evaluate the performance of the method for multi facies estimation with favorable outcomes. Additional examples with more complex channel features were used to evaluate the performance of the proposed approach under a less representative and completely inconsistent prior TI. When the true model contained features different from those commonly encountered in the TI, reasonable probability map updates were obtained. However, the method could not reproduce the correct facies orientation and structure when the TI was completely inconsistent with the true model. These examples highlight the importance of accounting for the uncertainty in the TI when solving the related MPS inverse problems. Furthermore, it is worth noting that, like all other inversion techniques, the proposed method relies on the information content in the data to resolve the features in the true solution. In cases where data resolution does not provide enough sensitivity to existing

features in the true model, the inversion approach would not be able to adequately constrain the simulated facies, even if a consistent training image was used.

A critical issue in the estimation of facies distribution is that of maintaining a balance between the prior information (such as a TI) and the observations. On one hand, the prior model plays a central role in the stabilization of history matching inverse problems that are known to be intrinsically ill-posed. On the other hand, excessive emphasis on a prior model can become a serious liability when the prior uncertainty is not adequately represented. The experiments presented in this section suggest that with a reasonably good prior TI, even limited production data can produce many acceptable conditional facies samples using the proposed PCM approach. However, when prior TI is not representative of the true model and the uncertainty in it is disregarded, the performance of the data integration formulation is degraded. The examples in this section clearly show that scattered flow data is not hard to match, even when prior TIs that are inconsistent with the true model are use. While this outcome is certainly not unique to the facies estimation framework discussed in here, it highlights the importance of a reliable TI for history matching problems, an issue that is at the heart of multipoint statistical simulation techniques.

In order to extend the application of PCM beyond the EnKF data integration technique, in the next section, an adaptive strategy is proposed through a data feedback mechanism based on the predictive performance of the past realizations (i.e., sampling history) to improve the conditioning efficiency. The method is presented as a stochastic optimization algorithm and extended to the case where multiple training images are

proposed as alternative plausible interpretation of the geologic scenarios for a given formation. Several numerical experiments from fluvial channel formations are used to demonstrate its applicability and performance in relation to non-adaptive conditioning techniques.

3. CONDITIONAL SAMPLING WITH PCM

3.1. Summary

Multiple-point statistics (MPS) provides a flexible grid-based approach for simulating complex geologic patterns with high-order statistics from a conceptual prior model known as a training image (TI). While MPS is quite powerful for describing complex geologic facies connectivity, the nonlinear and complex relation between the flow data and facies distribution makes flow data conditioning quite challenging. Here, an adaptive flow-conditioning method is proposed, through a data feedback mechanism, to simulate facies models from a prior TI. The adaptive conditioning uses a stochastic optimization algorithm with an initial exploration stage to identify the promising regions of the search space, followed by a more focused search in those regions in the second stage. To guide the search strategy, a facies probability map is constructed that summarizes the common features of the accepted models in previous iterations. The probability map provides additional conditioning information about facies occurrence in each grid block and is incorporated into the MPS algorithm as soft data to guide the facies simulation to generate a new candidate model. As the optimization iterations progress, the initial probability map is gradually updated based on the information about the facies distribution in the most recently accepted iteration. This conditioning process can be viewed as a stochastic optimization approach with memory where the new models are proposed based on the history of the successful iterations (with some randomness introduced during the MPS simulation process). The application of this adaptive

conditioning approach is extended to the case where multiple training images are proposed as alternative plausible interpretation of the geologic scenarios for a given formation. The advantages and limitations of the proposed adaptive conditioning scheme are discussed. Numerical experiments from fluvial channel formations are used to demonstrate its applicability and performance in relation to non-adaptive conditioning techniques.

3.2. Data Conditioning as an Optimization Problem

Subsurface environments consist of geological formations with complex spatial variability. Fluid flow in the heterogeneous porous media is controlled by the connectivity of the extreme (high/low) rock flow properties that are created by complex geological structures and lead to preferential flow paths/barriers. In many geological settings, the distribution of flow-related properties is primarily controlled by the location and distribution of distinct geologic facies. These properties typically exhibit sharp contrasts across facies boundaries (Weber, 1990). Hence, the fluid flow in such environments is strongly constrained by the connectivity and geometric attributes of the facies. Under these circumstances, the within-facies variability is of second order importance and the primary focus is on characterization of facies distributions, their spatial connectivity, and the associated uncertainties.

Multiple-point statistics (MPS), presented in detail in Subsection 1.3, was introduced as a flexible grid-based alternative method for simulation of complex geologic patterns. Conditioning the MPS simulation of facies to reproduce nonlinear

flow data poses a challenging inverse problem. Caers (2002) introduced the probability perturbation method (PPM), which is an iterative sampling approach where new proposal realizations are generated by parameterizing and perturbing the random seed number. The process is repeated until the generated realization is able to reproduce the observed data (within a specified error threshold) or the maximum number of iterations is reached (Caers, 2002; Hoffman and Caers, 2005; Kashib and Srinivasan, 2006). Alcolea and Renard (2010) present a block moving window algorithm for conditioning MPS simulations based on the block Markov Chain Monte Carlo method and use their method for integrating the pressure and connectivity data (Fu and Gomez-Hernandez, 2009). Hasen et al. (2008, 2012) propose a sequential Gibbs sampling approach that allows the Metropolis algorithm to efficiently incorporate complex priors into the solution of an inverse problem, which can be applied to the case where no closed form description of the prior exists. Mariethoz et al. (2010a) also discuss a similar approach, called the iterative spatial resampling method, to guide the simulation of new realizations to honor the observed data. Zhou et al. (2012) proposed a pattern searching inverse approach to infer static and dynamic state variables in an MPS approach using an ensemble of conductivity realizations and their respective heads as training data.

Monte-Carlo-based nonlinear inverse modeling methods pose the flow conditioning problem as either an iterative sampling problem or a post-simulation model updating step. The former approach attempts to sample the posterior distribution from the description of an explicit or implicit prior distribution and a likelihood model. Examples of this approach are Markov-Chain Monte-Carlo methods (Tarantola, 2005;

Gilks et al., 1996) and Particle Filters (Ristic et al., 2004), which are considered as theoretically rigorous but computationally prohibitive for large-scale problems.

Theoretically rigorous conditional sampling methods such as the particle filter and the Markov-Chain Monte Carlo (MCMC) methods have been found to be of limited practical appeal due to their prohibitive computational requirements when used with large-scale numerical simulation of coupled nonlinear differential flow and transport equations. The main computational burden of these Bayesian conditioning methods lies in the exceedingly large number of flow simulations that are needed for quantifying the likelihood function for a large number of model realizations. Conditional simulation in the MPS context is not exempt from the exceedingly high computational cost associated with the large number of flow simulations that must be performed. Thus, developing efficient conditional sampling techniques is critical for practical application of the MPS simulation.

In this section, adaptive conditional facies simulation is developed to speed up conditional simulation from a TI and the soft data conditioning approach in PCM (see Section 2 for more detail) is adopted. A main difference in the proposed approach with PCM is that, instead of using EnKF to generate the probability map, the facies probability maps is generated based on the common geologic patterns in a collection of accepted (conditioned) realizations. In the adapting conditional simulation approach, the generated realizations at early iteration have a relatively low chance of satisfying the observed data. However, as the iterations of the algorithm progresses the generated realizations become increasingly more likely to reproduce the observed flow data (within

the limits of observation noise), a property that significantly reduces the computational cost of facies conditioning. The algorithm begins by initially using a non-informative facies probability map to sequentially sample from a given TI. Each generated realization is used to predict the observed data and is either accepted or rejected based on a pre-defined probabilistic data mismatch threshold. As the sampling progresses, the accepted realizations and their statistical information are stored and used to update the spatial facies probability map to guide the simulation of new realizations. By using the updated facies probability map, this adaptive approach provides a feedback mechanism to indirectly incorporate the observed data into the sampling process and improves the computational cost of conditional MPS facies simulation. The application of the proposed method is extended to conditional facies simulation from multiple TIs that reflect the uncertainty in the geologic connectivity scenario. In the remainder of the section, the proposed method is presented in detail and several numerical experiments are used to discuss its performance and important properties.

3.3. Implementation

In the proposed method, using a feedback mechanism from the successful optimization iterations of the conditioning process, a facies probability map that contains information about the spatial distribution of discrete facies is inferred. This facies probability map is used as soft data (similar to the probability conditioning method) to guide conditional MPS simulation of facies within the *SNESIM* algorithm. As such, the facies probability map indirectly incorporates information about the flow data into the conditional MPS

simulation. In the following section, a simple non-adaptive formulation is presented followed by the implementation of the proposed adaptive approach.

3.3.1. Non-Adaptive Conditioning

To present the nonlinear inverse modeling for the flow and transport in porous media, y_{obs} represents the observed flow variables (rates or pressures). The predicted form of the observed quantities is denoted as $y = g(x)$, which characterizes a nonlinear mapping from model parameters to observed data. A deterministic inverse modeling formulation can be obtained by attempting to minimize the misfit between the predicted and observed response of the system. Assuming that the flow data is corrupted by a Gaussian noise with zero-mean and covariance C_{yy}, that is $y = g(x) + \varepsilon$, where $\varepsilon \sim G\left(0, C_{yy}\right)$, a common choice for the cost function is the weighted second norm of the data misfit function $\left\| W^{\frac{1}{2}}(g(x) - y_{obs}) \right\|_2^2$, which can be expressed as the following minimization problem

$$\min J(x) = (g(x) - y_{obs})^T C_{yy}^{-1} (g(x) - y_{obs}) \quad s.t. \;\; x \in TI \tag{3.1}$$

where the weighting matrix W is assumed to be the inverse of the observation error covariance matrix, that is $W = C_{yy}^{-1}$. Acceptable solutions of the above optimization problem constitute conditional realizations from the given TI. Given the discrete and non-differentiable nature of the facies as inversion parameters, non-gradient-based optimization methods are more appropriate for finding a solution. In particular, stochastic optimization algorithms do not require gradient information and can be suitable solution approach. A major complication in the above optimization problem is

the TI constraint $x \in TI$ which is used to ensure that solution must honor the higher order statistical patterns in the TI. Given this constraint, a simple solution strategy is to generate unconditional random realizations from the TI and examine their fitness by evaluating the above cost function. An example of such algorithm is given in **Table 3.1**.

Table 3.1 – Non-adaptive conditioning algorithm

- Initialize with a random realization $\mathbf{x}_0$ with objective function value $J(\mathbf{x}_0) = J_0$
- $\mathbf{x}_{\text{best}} = \mathbf{x}_0$ and $J_{\text{best}} = J_0$
- While $i < N_{max}$ and $J(\mathbf{x}_i) > \epsilon^2_{\text{obs}}$
    - Draw a random unconditional realization $\mathbf{x}_i$ and compute $J(\mathbf{x}_i) = J_i$
    - Perform acceptance test with a specified criteria, e.g. min $\left[1, \frac{\pi_i}{\pi_{i-1}}\right]$
        - If $\mathbf{x}_i$ is accepted, then $i = i + 1$,
        - If $J(\mathbf{x}_i) < J_{\text{best}} =$, then $\mathbf{x}_{\text{best}} = \mathbf{x}_i$ and $J_{\text{best}} = J(\mathbf{x}_i)$
        - $i = i + 1$
- Return $\mathbf{x}_{\text{best}}$

For acceptance criteria, the typical Gaussian likelihood ration functions π_i and π_{i-1} for iterates x_i and x_{i-1} are,

$$\pi(x_i) \propto exp\left(-\frac{1}{2}J(x_i)\right) \propto exp\left(-\frac{1}{2}(g(x_i) - y_{obs})^T C_{yy}^{-1}(g(x_i) - y_{obs})\right) \qquad (3.2)$$

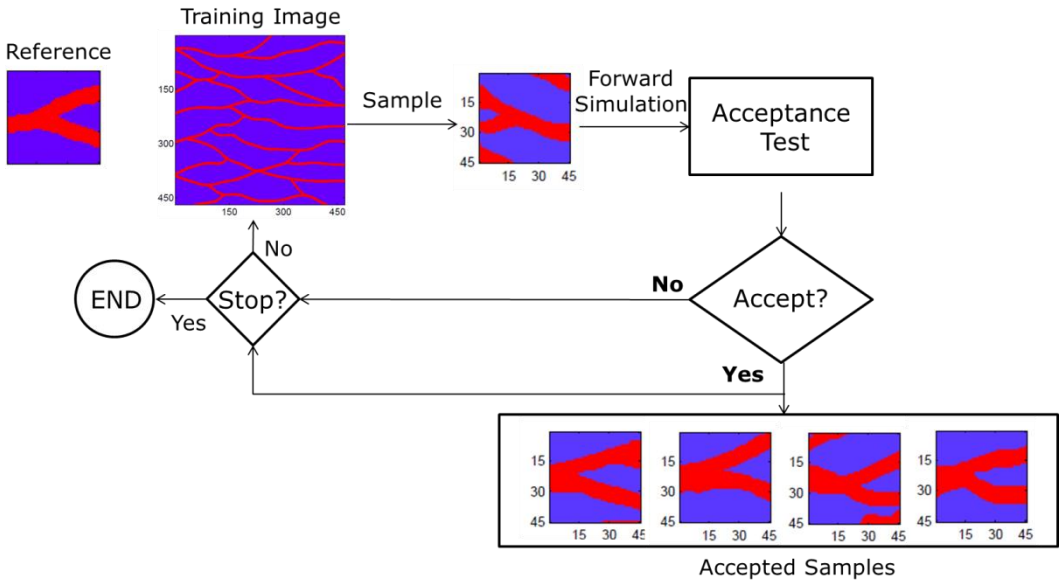
and use the likelihood ratios to accept or reject an iterate. Hence the ratio of the likelihood functions π_i and π_{i-1} for realizations x_i and x_{i-1} can be expressed as

$$\frac{\pi_i}{\pi_{i-1}} = \frac{\pi(x_i)}{\pi(x_{i-1})} = \frac{exp\left(-\frac{1}{2}J(x_i)\right)}{exp\left(-\frac{1}{2}J(x_{i-1})\right)} = \frac{exp\left(-\frac{1}{2}(g(x_i) - y_{obs})^T C_{yy}^{-1}(g(x_i) - y_{obs})\right)}{exp\left(-\frac{1}{2}(g(x_{i-1}) - y_{obs})^T C_{yy}^{-1}(g(x_{i-1}) - y_{obs})\right)} \qquad (3.3)$$

With this definition, the acceptance criterion is $min[1, \frac{\pi_i}{\pi_{i-1}}]$, which means that a proposed iterate x_i with probability 1 is accepted if it provides a better data match than the previous iterate x_{i-1} or with probability $\frac{\pi_i}{\pi_{i-1}} < 1$ if the data match is not improved by the current. **Fig. 3.1(a)** depicts the non-adaptive facies conditioning workflow with a TI representing a simple fluvial channel system that will be discussed in the next section.

A major limitation of the above algorithm is its computational inefficiency. In general, when the computation of the objective function is fast, one could set the number of iterations (N_{max}) to be very large and randomly generate facies models to implement the steps in Table 3.1 to explore the search space and identify the model with the best data match. However, given the complexity of facies distribution in large models and the high computational cost of forward model simulation to evaluate the objection function for each proposed model (iterate), the above algorithm, while simple, offers an extremely inefficient and impractical procedure for generating conditional facies models. In general, stochastic optimization methods involve a large number of iterations to explore the search space to seek the global optimizer. They typically start with a broad search (exploration) of the parameter space to locate promising regions and then narrow the search to the proximity of the identified neighborhoods to find the optimal solution. The simple algorithm described in Table 3.1 does not follow this two-stage approach, which increases its computational complexity. An adaptive algorithm is proposed to improve the computational efficiency of the algorithm.

(a) Non-Adaptive conditioning method



(b) Adaptive conditioning method

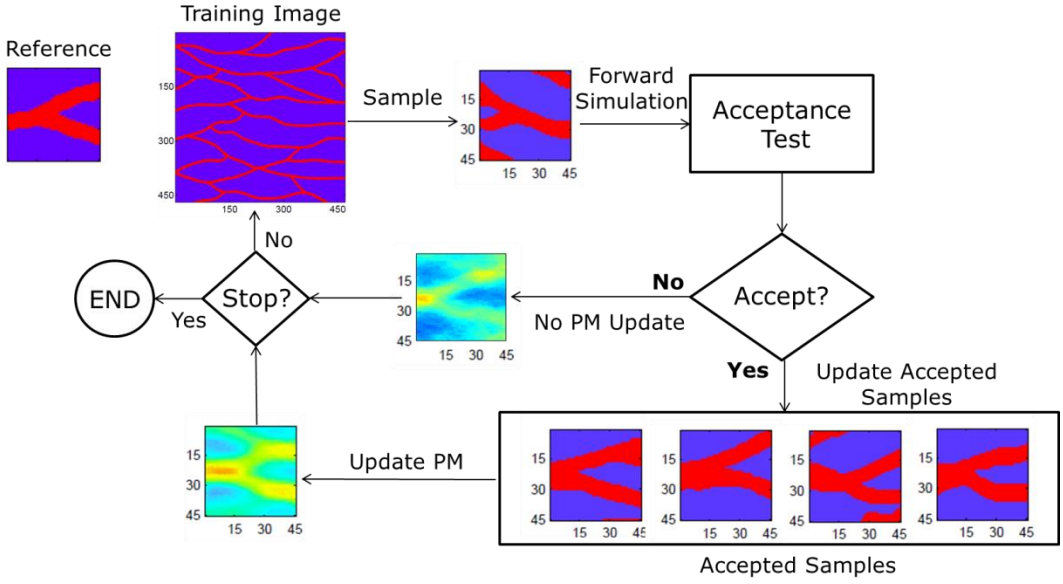


Fig. 3.1 – Flowchart diagram showing the implementation steps of (a) non-adaptive and (b) adaptive conditioning methods.

3.3.2. Adaptive Conditioning with Facies Probability Maps

To improve the efficiency of the proposed method, an adaptive strategy is proposed in which initially the algorithm explores the search space following the non-adaptive procedure in Table 3.1. This step is used to explore different regions of the search space and identify possible local solutions. The accepted models (iterates) are used to generate a facies probability map that will be included in the MPS simulation to guide the generation of a new proposal facies model. The intended purpose of the facies probability map is to improve the conditioning step of the algorithm by incorporating, probabilistically, the common connectivity patterns among the accepted models from previous iterations in the MPS simulation.

To develop an adaptive approach, the accepted models are stored and used to improve the performance of the conditional facies simulation from the TI. Specifically, the algorithm uses the statistical information about the accepted models to generate a facies probability map. The facies probability map is then incorporated into the SNESIM algorithm as soft data to simulate facies models with higher acceptance probability. In this case, the conditioning process is not independent and has a memory. The facies probability map is initially uninformative and is updated after each successful iteration to reflect the new information provided by the accepted model. The updates applied to the facies probability map are very gradual to ensure that: (1) initially, when the probability map is uninformative, the facies models drawn mostly follow the variability in the TI with little constraining from the TI; and (2) at later stages, when the probability

map is more informative, relatively smaller update steps are introduced to primarily explore the neighborhood around the probability map.

The implementation is begun by generating the first facies model with an uninformative constant probability value proportional to the ratio of different facies types (the marginal distribution of the facies). This will ensure consistency with the TI prior to data integration. After assigning permeability values to the facies, the reservoir model is used to predict the observations that are needed to calculate the fitness function for the new facies model. If a new model is accepted, it is used to update the facies probability map. The following equation is adopted to update the probability map after each successful conditioning iteration:

$$PM_k = \begin{cases} \left(1 - \frac{k}{N_{bi}}\right) \times PM_{init} + \frac{k}{N_{bi}} \times \sum_{j=1}^{N_{acpt}} w_j \mathbf{x}_j^{acpt} & k \leq N_{bi} \\ \sum_{j=1}^{N_{acpt}} w_j \mathbf{x}_j^{acpt} & k > N_{bi} \end{cases} \tag{3.4}$$

where k is the iteration number, N_{acpt} is the number of accepted iterations, w is the weight given to accepted model at each iteration (proportional to its predictive performance), N_{bi} is the number of burn-in iterations in which the effect of the initial (uninformative) probability map is more dominant for a broader exploration of the search space. High N_{bi} values correspond to smaller corrections to the initial probability map (longer memory for the initial probability map) and extended exploration time; in the limit as $N_{bi} \to \infty$ the behavior of the adaptive algorithm becomes similar to the non-adaptive case. On the other hand, lower values for N_{bi} result in strong (premature)

updates to the facies probability map, which is not desirable at the initial stage of the algorithm. $N_{bi} = 100$ is used for the numerical experiments here. The implementation steps for the adaptive algorithm are summarized in **Table 3.2**. While in this implementation the maximum number of iterations and the value of the objective function are the stopping criteria of the optimization, the facies probability map may also be used to define a stopping criterion. For example, one could stop the optimization if the changes in the probability map in the last k iterations falls below a predefined threshold.

Table 3.2 – Adaptive conditioning algorithm

- Initialize with a realization x_0 (and $J(x_0) = J_0$) and N_{acpt}
- Generate an uninformative probability map PM_{init}
- Assign $x_{best} = x_0$ and $J_{best} = J_0$ and $PM_0 = PM_{init}$
- While $i < N_{max}$ and $J(x_i) > \epsilon_{obs}^2$
    - Using PM_k to draw a conditional realization x_i and compute $J(x_i)$ and π_i
    - Perform acceptance test with a specified criteria, e.g. min $\left[1, \frac{\pi_i}{\pi_{i-1}}\right]$
        - If x_i is accepted, then
        - update PM_k using Eq. (3.4)
        - $k = k + 1$
    - If $J(x_i) < J_{best}$ then
        - $x_{best} = x_i$ and $J_{best} = J(x_i)$
    - $i = i + 1$
- Return x_{best}

The adaptive algorithm in Table 3.2 bears similarity to the simulated annealing algorithm, with the difference that instead of the temperature parameter that controls the temperature decrease rate (transition from exploration to focused search), it uses the

probability map to control the search. However, it is important to note that application of the stochastic search methods such as simulated annealing to MPS-based flow data conditioning is complicated by generating controlled perturbations that are consistent with TI statistics. The probability perturbation method (Caers, 2002) automatically satisfies the TI consistency constraint by introducing the perturbation to the parameterization of the random seed. The proposed adaptive approach based on the facies probability map can introduce perturbation that are consistent with the TI and, additionally, provide a flow-data feedback mechanism to guide the conditional MPS simulation outcome. The feedback mechanism through facies probability map is general and can be in included in other conditioning techniques such as the PPM. In addition, in the current implementation the marginal distribution of facies for each gridblock from the accepted iterations to is used define the conditioning facies probability map. Alternative implementations can be developed to incorporate higher-order statistical patterns from the accepted iterations into the conditioning process.

Fig. 3.1(b) displays the adaptive conditioning workflow in contrast to the non-adaptive cases shown in Fig. 3.1(a). The adaptation may be interpreted as a proposed search direction that is derived from the performance of the accepted models in previous iterations. The resampling or re-simulation approach (Hasen et al., 2008, 2012; Mariethoz et al. 2010a) take advantage of the accepted samples in generating a new proposal model (perturbation). In my case, the conditioning is guided by a probability map that represents the statistical information about the facies distribution from previously accepted iterates. Finally, after convergence when the changes to the

probability map becomes minimal the adaptive approach may be used to generate several facies models with the final probability map to explore the local variability around the obtained solution in the search space.

3.3.3. Adaptive Conditional Facies Simulation from Multiple TIs

The optimization algorithms described above can be extended to the case where multiple TIs are used to account for the uncertainty in the form and structure of facies connectivity that are presented as plausible geologic scenarios. The MPS simulation offers a pattern imitating replication of the connectivity information embedded in a TI. As such, it is able to preserve the higher order statistics of the TI patterns, which is very desirable if the TI adequately represents the reality. When the TI (i.e., the conceptual prior connectivity model) is not correctly specified, a persistent bias is introduced in the simulation results. To account for the uncertainty in the TI, one practical approach is to use multiple TIs that provide alternative interpretations, for example by different geologists, of the available data and geologic knowledge. The uncertainty in the TI may also be modeled through stochastic treatment of the input parameters in the process-based modeling. In generating conditional samples from multiple TIs one is confronted with two problems: 1) identifying the relative importance (weight) of each TI and 2) conditioning on flow data. The resulting problem is clearly more complicated than conditional sampling from a single TI. Since the flow-data feedback mechanism in our adaptive conditioning approach provides information about the performance of each TI in matching the flow data, it can be used to address the two problems at once.

Here, the adaptive conditioning approach is extended to flow data conditioning under TI uncertainty. To this end, the algorithm is modified to generate a conditional model from multiple TIs each with a weight proportional to its predictive performance. Specifically, during the exploration time, each TI is given an equal weight (that sum to one) and the adaptive algorithm is initially run independently for each TI. The accepted iteration from each TI will be used to update the probability map for the corresponding TI. This results in a separate probability map for each TI. After the initial stage, the accepted iterations from all TIs are given both a normalized weight and a TI flag that shows from which TI they are simulated. The normalized weight w_j for each accepted model x_j^{acpt} is proportional to the function $\pi(x_j)$. The TI weights are calculated as the sum of the weights of the accepted models drawn from the corresponding TI. The accepted models from each TI are then used in Eq. (3.4) to generate a corresponding probability map for that TI. The next model is then proposed by first probabilistically selecting a TI (based on TI weights) and then simulating a facies model using the corresponding probability map. As the iterations proceed and the TI weights are updated, more (fewer) models are proposed (drawn) from the TIs with the larger (smaller) weights. In the next section, the application of the above adaptive conditioning algorithms with three numerical experiments is demonstrated.

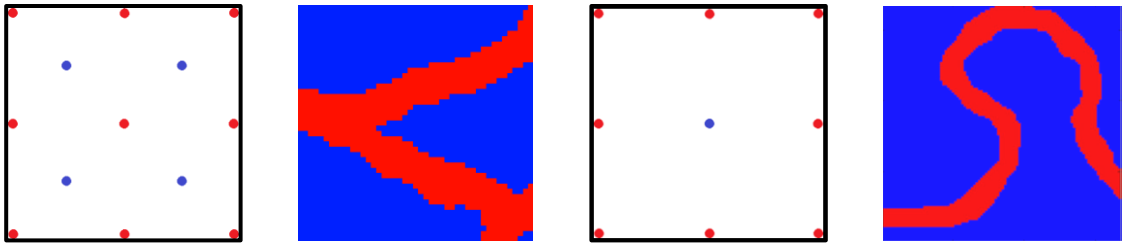


Fig. 3.2 – Well configuration and true synthetic models for the two water flooding examples: (a) 13-spot well configuration in a 45×45 two-dimensional reservoir and (b) reference facies model for Example 3.1; (c) 9-spot well configuration in a 100×100 two-dimensional reservoir and (d) reference meandering facies model for Example 3.2 and Example 3.3.

3.4. Applications

To investigate the efficiency of the proposed adaptive conditional simulation approach, the non-adaptive algorithm is applied to two water flooding examples to estimate the distribution of channel facies from production measurements. In the third example, three structurally different TIs are provided to account for the uncertainty in the prior channel connectivity model. For each example, the non-adaptive algorithm is used followed by the adaptive conditioning under identical conditions. This section is concluded with a discussion about the results and the properties of the proposed adaptive algorithm.

3.4.1. Example 3.1: Straight Channels

3.4.1.1. Experimental Setup

In the first example, a 13-spot reservoir configuration consisting of four water injection wells and nine oil production wells is considered. The objective is to infer the two-

dimensional facies distributions from the production observations in these wells. The facies permeability values are known while their spatial distribution is unknown. The channel and background facies permeability values are $200mD$ and $10mD$, respectively.

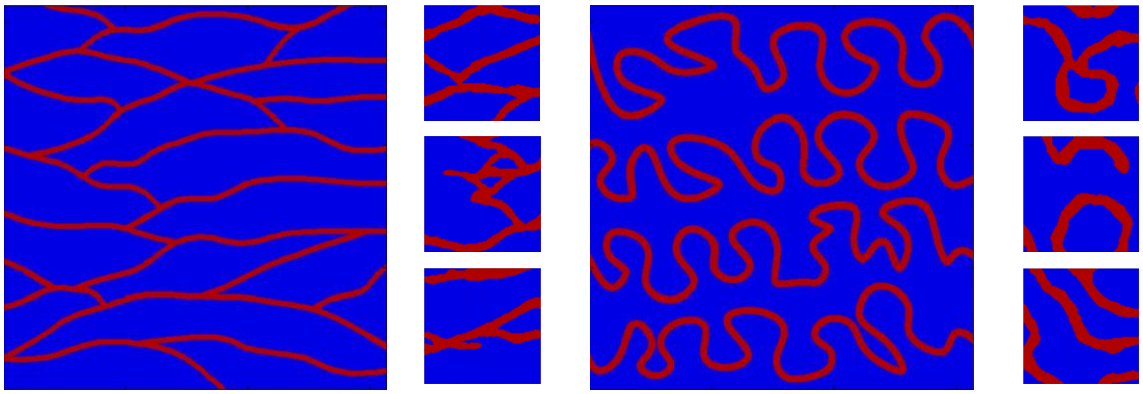


Fig. 3.3 – The training images and sample realizations for Examples 3.1 and 3.2; (a)-(b) The TI and sample facies realizations for Example 3.1; (c)-(d) the TI and sample facies realizations for Example 3.2.

The initial oil saturation and pressure are 0.90 and $3000psi$, respectively, and are distributed uniformly throughout the reservoir. The injection wells operate with the rate of $300STBD$ during the twelve years of simulation. The production wells operate under a constant pressure of $1000psi$. Under these settings, the pressure at the injection wells as well as the water and oil rates at the production wells were measured every year and used for facies conditioning. The well configuration and reference permeability map for Example 3.1 are illustrated in **Fig. 3.2(a)** and **3.2(b)**, respectively. For the first

example, the performance of the method is investigated under different data noise levels (signal-to-noise ratios) as discussed in the next section.

3.4.1.2. Non-Adaptive Conditioning

First, the non-adaptive algorithm is applied for the conditional facies sampling. The synthetic true facies model is derived from the TI in **Fig. 3.3(a)** and contains two intersecting channels. Three unconditional facies realizations from the same TI are depicted in Fig. 3.3(b). In this experiment, the optimization algorithm is run for 10,000 iterations. The best models at select iterations are shown in **Fig. 3.4(a)**. The first to the last columns of Fig. 3.4(a) show the best reservoir model (with the lowest data misfit) obtained at iterations 100, 500, 1000, 5000, 10000, respectively. Approximately 38.31% (3831 out of 10,000) of the simulated realizations in this example were accepted. Note that acceptance of an iteration only implies that the current iterate improves the best available solution; clearly, initially the algorithm has a high acceptance rate since the initial model is generally far from the solution. This can be observed from the results in Fig. 3.4(a) that show the best model up to 1000^{th} iteration is still far from the true model. The best reservoir model after 5000 iterations (Fourth column in Fig. 3.4(a)) begins to resemble the true synthetic model.

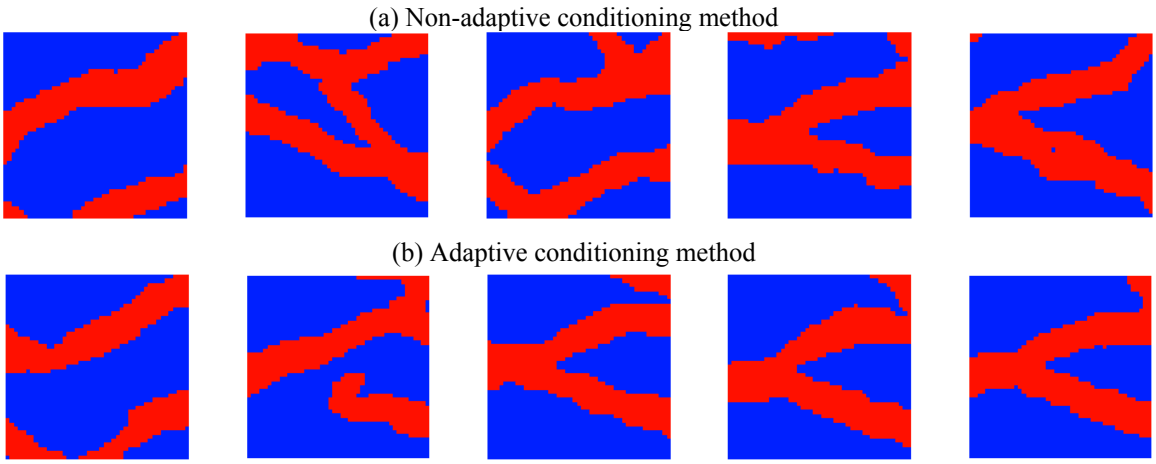


Fig. 3.4 – The best facies model (solution) after 100 (1st column), 500 (2nd column), 1000 (3rd column), 5000 (4th column), and 10000 (last column) iterations, for the non-adaptive (a) and adaptive (b) conditioning methods in Example 3.1.

3.4.1.3. Adaptive Conditioning

The results of conditional facies simulation using the proposed adaptive conditioning approach for the first example are illustrated in Fig. 3.4(b). The columns in Fig. 3.4(b) show, from left to right, the best reservoir model (with the lowest data misfit) obtained at iterations 100, 500, 1000, 5000, 10000, respectively. The total number of accepted iterations in this case is 4809 (48.09%), which is 10% higher than that of the non-adaptive case. While this shows an improvement over the non-adaptive approach, it does not highlight the convergence behavior of the adaptive method. The behavior in Fig. 3.4(b) is clearly different from that in Fig. 3.4(a). In Fig. 3.4(b), the best reservoir

models that are generated after only 500 iterations have similar connectivity features to those in the true model, which indicates a much faster convergence behaviors.

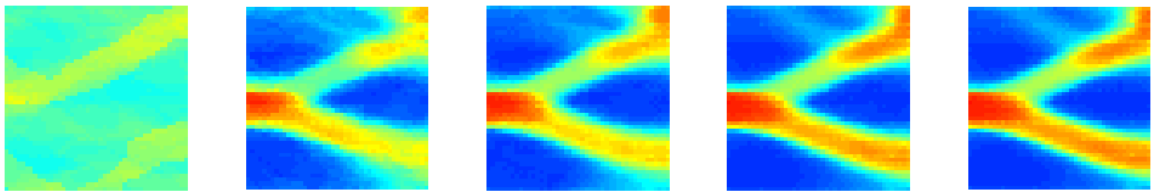


Fig. 3.5 – The updated probability map after 100 (1st column), 500 (2nd column), 1000 (3rd column), 5000 (4th column), and 10000 (last column) iterations, for the adaptive conditioning method in Example 3.1.

Fig. 3.5 shows the evolution of the probability map with increasing number of iterations for the adaptive conditioning case. The main channel features in the reference model are clearly distinguishable in the probability map. This implies that the facies models generated with the adaptive approach gradually improve with increasing number of iterations. Hence, the simulated models tend to explore the facies distribution at the proximity of the connectivity pattern identified in the TI. These results clearly show the adaptive nature of the algorithm and the resulting increase in the efficiency of the conditioning approach. **Fig. 3.6** shows the flow prediction performance of the two methods. The plots in Fig. 3.6 display the RMSE of the best reservoir model at different iterations for the non-adaptive (red curve) and adaptive (blue curve) algorithms. The lowest RMSE for adaptive and non-adaptive algorithms do not change significantly after

682 and 2461 iterations, respectively. Furthermore, the data match quality of the conditional model obtained from the adaptive approach is superior to the solution obtained from the non-adaptive case, implying that, in this example, the adaptive algorithm finds a better solution in less iteration.

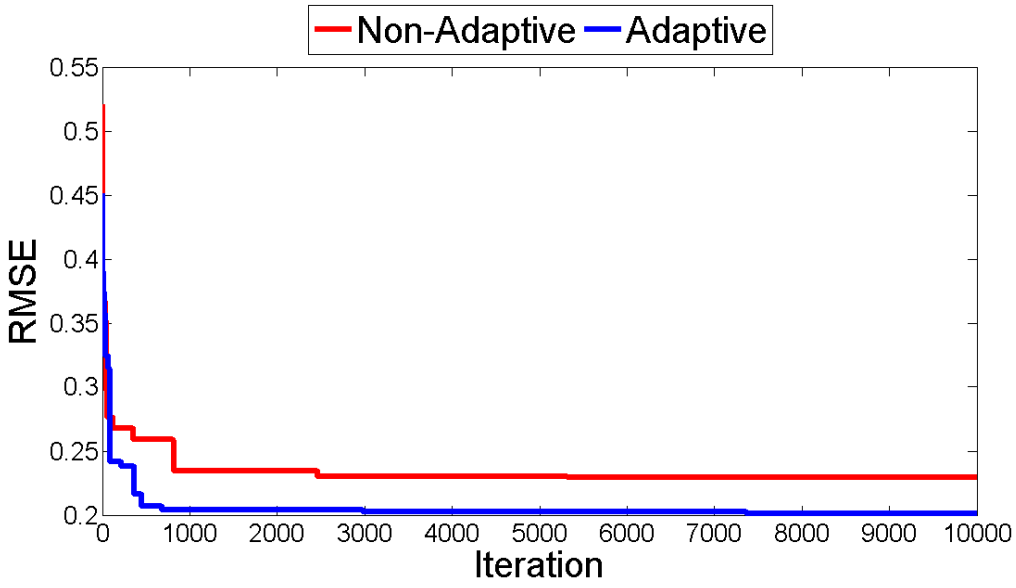
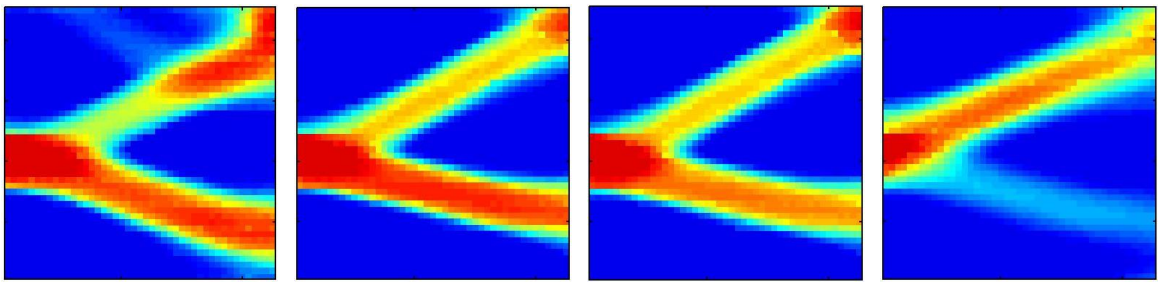


Fig. 3.6 – RMSE of the best facies model (solution) for the non-adaptive (red line) and adaptive (blue line) conditioning methods in Example 3.1.

Also the performance of the adaptive algorithm is investigated under observation noise. **Fig. 3.7** shows the final probability map that is obtained under different noise levels. The noise level is presented as the ratio (in percentage) of the zero-mean Gaussian noise standard deviation to the maximum range of variability in the

corresponding data values. While the conditioning performance degrades with increasing magnitude of the observation noise (decreasing signal-to-noise ratio), within a reasonable noise level, the results do not show significant deterioration. At a noise level of 30%, a major drop in the quality of the accepted models is observed where the lower channel begins to gradually disappear (Fig. 3.7(d)).



(a) (b) (c) (d)

Fig. 3.7 – Final probability maps in Example 3.1 with (a) 0% (original), (b) 10%, (c) 20%, and (d) 30% added observation noise. The noise standard deviation is reported as the percentage of the maximum range of variability in the observed data.

Fig. 3.8 displays the reservoir model with the lowest data misfit at iterations 100, 500, 1000, 5000, 10000 for noise levels of 0%, 10%, 20%, and 30%, respectively. The algorithm does not seem to show significant sensitivity to the noise, at least at the moderate noise levels that were considered in this study.

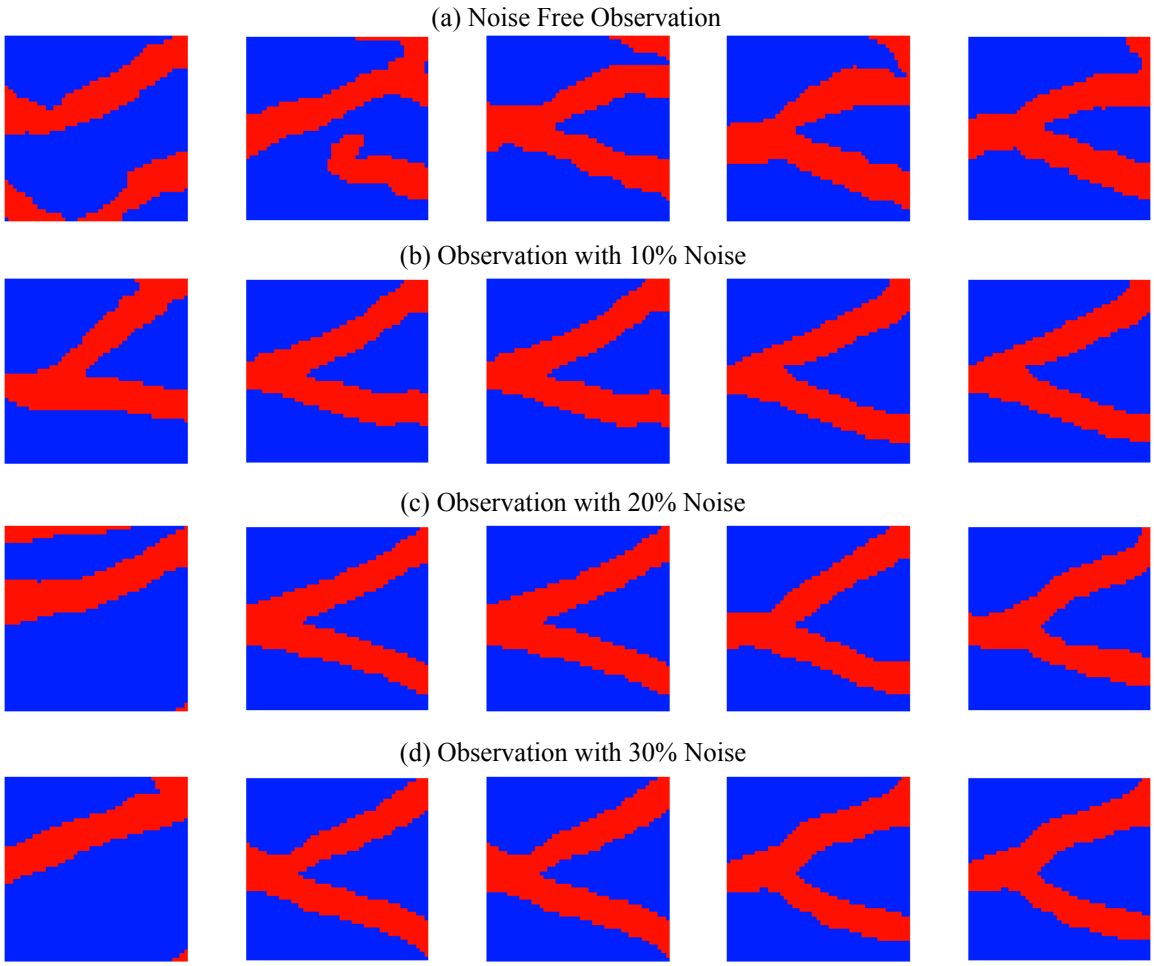


Fig. 3.8 – The best facies model (solution) after 100 (1st column), 500 (2nd column), 1000 (3rd column), 5000 (4th column), and 10000 (last column) iterations in Example 3.1: (a) no observation noise, (b) 10% observation noise, (c) 20% observation noise , and (d) 30% observation noise.

3.4.2. Example 3.2: Meandering Channels

3.4.2.1 Experimental Setup

In the second example, a 9-spot reservoir configuration consisting of one water injection well and eight oil production wells is adopted. Similar assumptions to those in Example

3.1, including the well controls and their specified operational values, the initial oil saturation and pressure in the field as well as the measured production quantities, are assumed. The measurements were obtained every 90 days. The channel and background facies permeability values are $200mD$ and $10mD$, respectively. The well configuration and the true synthetic permeability map are illustrated in Fig. 3.2(c) and 3.2(d), respectively.

3.4.2.2 Non-Adaptive Conditioning

The synthetic true permeability model in Example 3.2 is derived from the TI in Fig. 3.3(c). This model contains a meandering channel with an overall east-west orientation. Three unconditional reservoir models are depicted in Fig. 3.3(d). Only the injection well and one of the production wells (located in the north of the reservoir) are drilled in the channel section of the reservoir. The simulation results with the non-adaptive algorithm are shown in **Fig. 3.9(a)**. The columns, from left to right in Fig. 3.9(a) show the best reservoir model obtained at iterations 100, 500, 1000, 5000, 10000, respectively. While the obtained best model shows meandering patterns that are similar to those in the reference model (Fig. 3.2(d)), the feature in the lower part of the domain is not captured well.

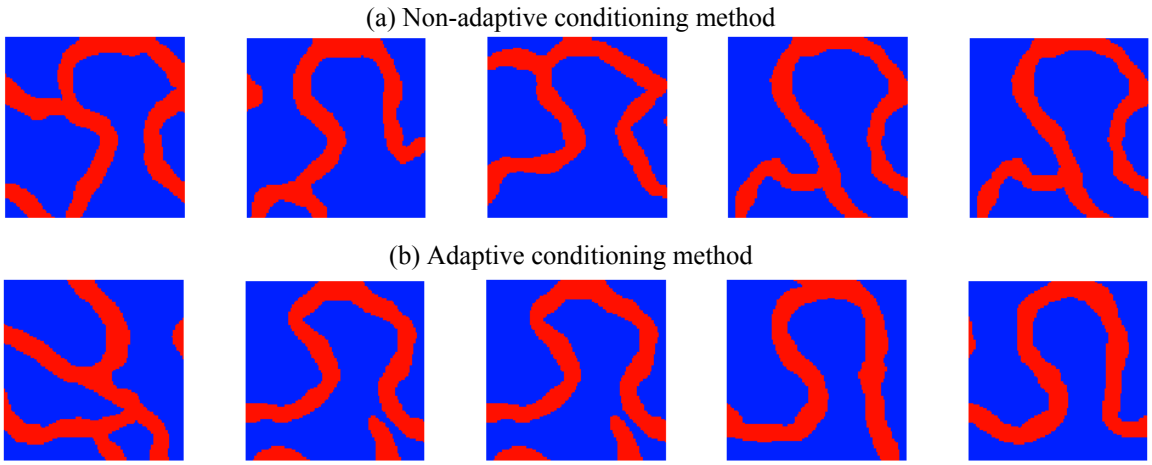


Fig. 3.9 – The best facies model (solution) for the non-adaptive (a) and adaptive (b) conditioning methods in Example 3.2 after 100 (1st column), 500 (2nd column), 1000 (3rd column), 5000 (4th column), and 10000 (last column) iterations.

3.4.2.3. Adaptive Conditioning

The results of facies conditioning for Example 3.2 using the proposed adaptive approach are shown in Figs. 3.9(b) and **3.10**. The columns, from left to right, in Fig. 3.9(a) show the best reservoir model obtained at iterations 100, 500, 1000, 5000, 10000, respectively. In this case, the final solution for the adaptive case is slightly better. Fig. 3.10 contains the evolution of the channel probability map at different optimization iterations for the adaptive implementation. The columns, from left to right, in Fig. 3.10 display the updated probability map at iterations 100, 500, 1000, 5000, 10000, respectively, which show that as the iterations proceed the probability map continuously improves. However, the improvement at later iterations is less pronounced. This can be partly attributed to the larger size of the domain (increased variability), the fewer number of wells (data) and

the complexity of the meandering channel facies. In general, these factors can make the feedback mechanism less effective. Nonetheless, the results show that the probability map provides important information about facies connectivity.

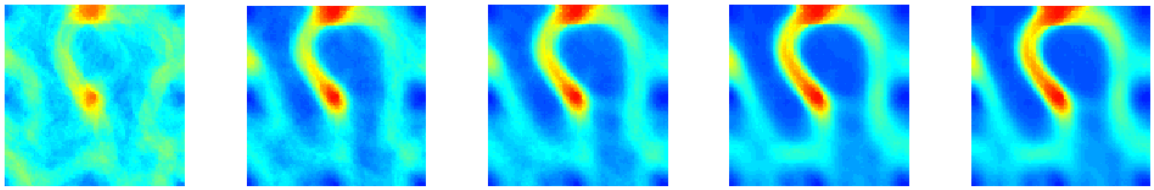


Fig. 3.10 – The updated probability map after 100 (1st column), 500 (2nd column), 1000 (3rd column), 5000 (4th column), and 10000 (last column) iterations, for the adaptive conditioning method in Example 3.2.

Finally, **Fig. 3.11** plots the evolution of the RMSE value for the best model with increasing number of iterations for both the non-adaptive (red line) and adaptive (blue line) algorithms in Example 3.2. The RMSE of the best model obtained from the adaptive algorithm is lower than the one obtained from the non-adaptive approach, which is consistent with the results in Example 3.1.

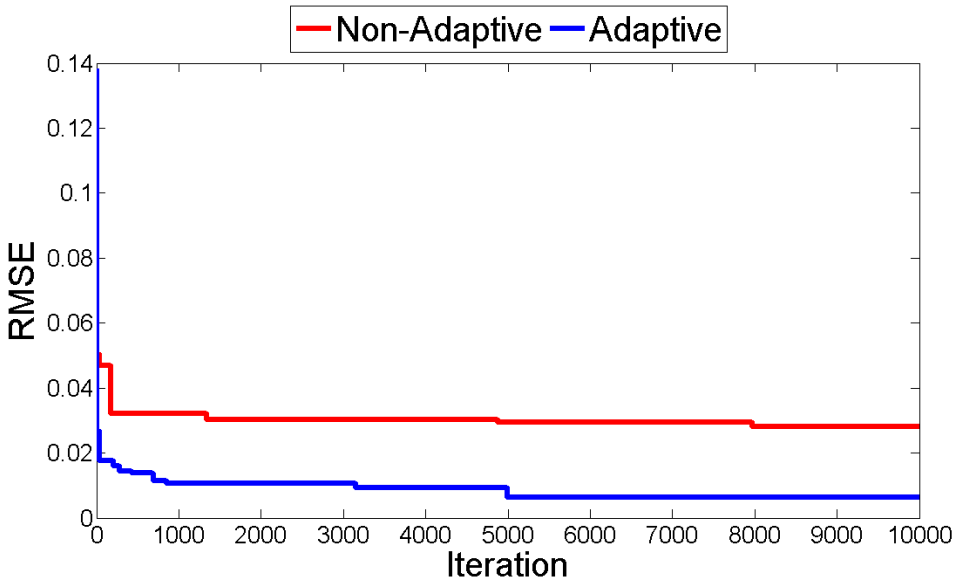


Fig. 3.11 – RMSE of the best facies model (solution) for the non-adaptive (red line) and adaptive (blue line) conditioning methods in Example 3.2.

3.4.3. Example 3.3: Multiple TIs

3.4.3.1 Experimental Setup

In the last example of this section, the application of the adaptive conditioning algorithm is considered when multiple TIs are used to represent the uncertainty in the prior model. The reservoir configuration and experimental conditions remain identical to those used in Example 3.2. However, as shown in **Fig. 3.12**, in this case we consider three alternative TIs with distinct connectivity patterns for describing the type of variability in the fluvial formation.

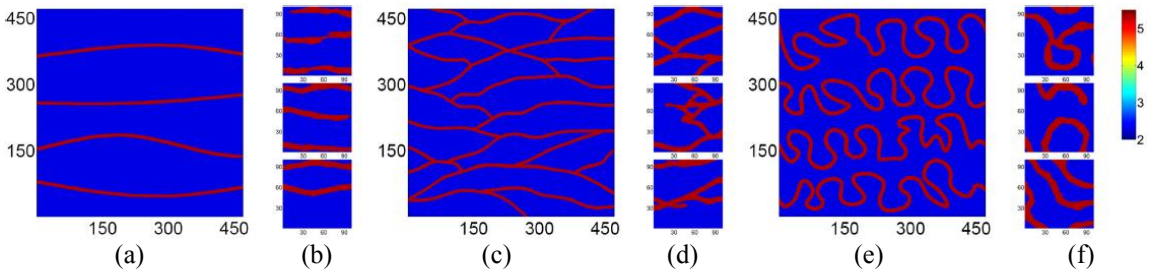
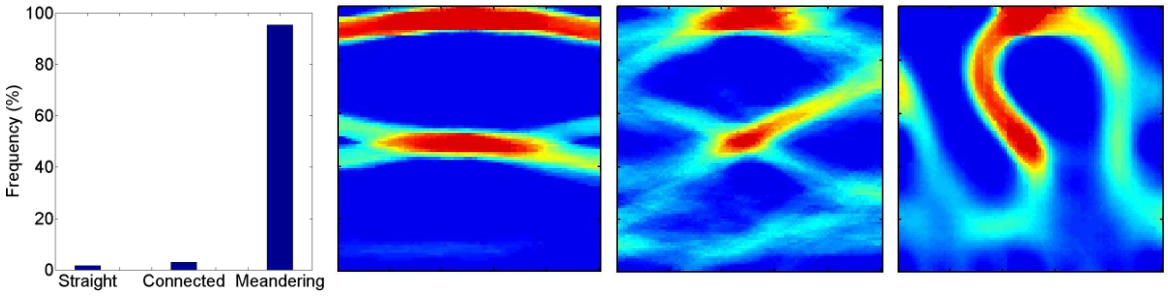


Fig. 3.12 – Three TIs with different structural connectivity each with three corresponding unconditional MPS realizations: (a) the TI with straight channels and (b) three corresponding realizations; (c) the TI with intersecting straight channels and (d) three corresponding MPS realizations; and (e) the TI with meandering channels and (f) three corresponding MPS realizations.

Figs. 3.12(a), 3.12(c), and 3.12(e) show the three TIs that exhibit straight left-to-right non-intersecting channels, predominantly left-to-right but intersecting channels, and meandering channels, respectively. Figs. 3.12(b), 3.12(d), and 3.12(f) show three facies models that are simulated from the corresponding TIs in Figs. 3.12(a), 3.12(c), and 3.12(e), respectively. The reference facies map belongs to the more complex meandering channel in Fig. 3.12(e). The reference model in this example is the same as that in Example 3.2 and belongs to the TI with the meandering channel. The objective is to implement the adaptive conditioning approach of Subsection 3.3.3 and evaluate its performance under TI uncertainty.



(a) (b) (c) (d)

Fig. 3.13 – The results for adaptive facies conditioning from multiple TIs: (a) TI weights (%) for each training image in Fig. 3.12; (b)-(d) show final probability map for the TIs with straight, intersecting and meandering channels, respectively.

3.4.3.2 Adaptive Conditioning

In this example, initially each TI was given an equal weight of 1/3. As the iterations proceed beyond the burn-in stage, the TI weights are updated based on the data match performance of the accepted models from each TI (as described in Subsection 3.3.3). **Fig. 3.13** summarizes the adaptive conditioning results. Fig. 3.13(a) displays the final TI weights in percentage while Figs. 3.13(b-d) show the probability map corresponding to each individual TI. The best solution at different stages (iterations 100, 500, 1000, 5000, and 10000) of the algorithm are shown in **Fig. 3.14** and suggest that in all cases the best samples belong to the TI with the meandering channel. Finally, the RMSE of the data match is shown is **Fig. 3.15**. While the RMSE tend to decrease even beyond 4000th iteration, most of reduction in RMSE occurs within the initial stages. It is clear from the results that the TI with meandering channels has the dominant contribution to the conditional sampling process. However, the contributions from other TIs are non-zero,

indicating that a number of facies models from other TI were accepted. However, the results indicate that these samples were accepted at the early stages where the best model is still not close to the reference model, and hence, the acceptance probability is high.

Fig. 3.14 – The best facies model (solution) after 100 (1st column), 500 (2nd column), 1000 (3rd column), 5000 (4th column), and 10000 (last column) iterations for Example 3.3 where adaptive facies conditioning from multiple TIs is used.

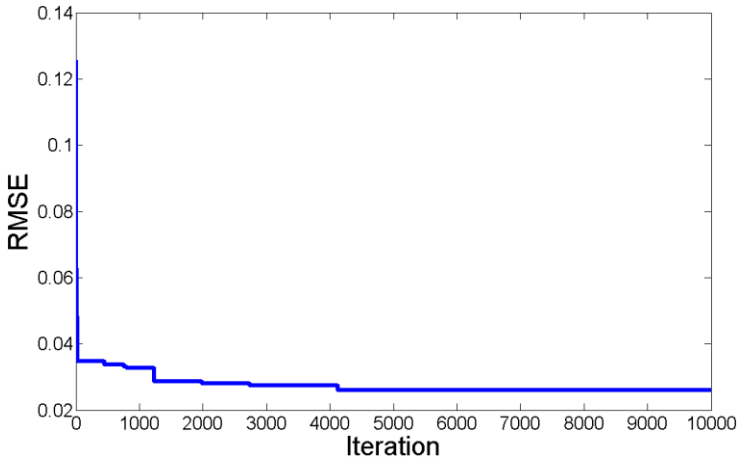


Fig. 3.15 – The RMSE plot for the best facies model (solution) for the adaptive conditioning method in Example 3.3.

While the facies distribution resulting from the inconsistent TIs cannot capture the correct connectivity in the reference model, they seem to identify the local variability around the wells in the reference model. In general, non-zero contribution from inconsistent TIs is possible and could be explained by the ill-posed nature of the inverse problem where alternative geologic scenarios may be found to reproduce the limited observed data (albeit with lower probability). In this particular example where the contrast between consistent and inconsistent TIs is quite noticeable, incorporating a data-feedback mechanism can effectively distinguish between consistent and inconsistent TIs.

3.5. Conclusions

A stochastic optimization framework is presented for adaptive conditioning of MPS-based facies simulation from a TI to nonlinear flow data. This was achieved by learning a facies probability map that describes the likelihood of facies occurrence in each grid block from accepted facies models in previous iterations. The optimization is initialized with an uninformative probability map and an unconditional facies model. The optimization algorithm consists of two stages: (i) a burn-in time in which the update to the probability map is minimal to explore the parameter space and (ii) a focused search stage where the promising regions of the parameter space are extensively explored to find the optimal solution. The probability map is gradually updated by incorporating the information about the facies distribution from the chain of accepted iterates. Since the accepted facies models improve the data match the dominant facies connectivity in them (common features among the accepted models) are likely to represent some of the

existing trends in the solution. Hence, a collective representation of the common features in the accepted models is used to construct a facies probability map as a data-feedback mechanism to guide and accelerate the conditioning process.

The implementation of the proposed approach is illustrated its performance using numerical water-flooding experiments with a single given TI and where multiple TIs were used to reflect the uncertainty in the geologic scenario. The proposed adaptive algorithm was able to effectively condition facies simulation from a given TI to the observed flow data. The feedback from the observed data into the simulation process was performed through updated facies probability maps and provided an effective mechanism to guide the simulation of the new facies models and improve the efficiency of the conditioning algorithm. The proposed adaptive approach generates more representative facies models in less iteration. Thus, compared to non-adaptive methods, adaptive algorithms based on probability conditioning are expected to provide significant computational saving in practical applications. It is important to note that the feedback mechanism used in the adaptive algorithm is general and can be included in existing conditioning algorithms. Also, the proposed adaptive algorithm is extended to applications with multiple TIs. Application to a synthetic example indicated that the adaptive algorithm can distinguish between consistent and inconsistent TIs and adjust the weight of each TI according to their contribution to explaining the observed data.

An important aspect of the proposed algorithm that requires further investigation is a sensitivity analysis with respect to different parameters of the algorithm and its ability to adequately explore the parameter space. While the probability map in these

examples converged to a pronounced connectivity features that were consistent with those in the reference model, this result cannot be generalized. In general, the adaptive algorithm results in a single probability map that may be sensitive to different initialization and/or random seed selection. The effect of multimodality (multiple local solutions) on the probability map is an important topic that was not considered in this research work. It is also important to allow for an extended exploration of the prior parameter space (the TI) especially in large scale problems with complex geologic connectivity where multiple local solutions with distinct connectivity patterns are likely to exist. Parallel implementation with different initialization may be useful for investigating the robustness of the algorithm against the initialization parameters (Romary, 2010). Furthermore, without proper implementation, the adaptation introduced by the feedback mechanism can limit the initial exploration of the parameter search space. This effect can be mitigated through several mechanisms such as dampening the update to the probability map by the accepted iterations, balancing the importance of the TI and the probability map, regionalizing the definition of the probability map to increase the diversity in areas with fewer or no observations, and adjusting the lower and upper bounds on the probability map to control its relative impact on facies simulation. Furthermore, implementation of the adaptive algorithm in parallel chains by initializing it from different points in the search space can provide additional insight into the convergence and robustness properties of the proposed method.

In the next section, a novel Bayesian mixture model for adaptive and efficient sampling of conditional facies from multiple uncertain TIs is developed. The posterior

distribution of facies is partitioned into individual conditional densities of the TIs and the corresponding mixture weights are estimated from the likelihood function for each TI. The suitability of the proposed Bayesian mixture-modeling approach is demonstrated using several numerical experiments in fluvial formations with uncertain orientation and structural connectivity.

4. MIXTURE MODEL PROBABILITY CONDITIONING METHOD (MMPCM)*

4.1. Summary

Multiple-point statistics (MPS) provides a systematic approach for pattern-based simulation of complex discrete geologic objects from a conceptual training image (TI) as prior model. The TI contains the general shape, geometry, and connectivity structures of complex patterns and encodes the related higher-order spatial statistics of the expected features. Conditioning MPS simulated facies on flow data poses a challenging nonlinear inverse problem for estimating discrete parameter fields. Additionally, the pattern-imitating nature of MPS simulation implies that the simulated facies inherit the spatial structure of the features in the TI. Since TIs are constructed from uncertain geologic information and imperfect assumptions, the resulting simulated facies may fail to predict the correct flow and transport behavior in the subsurface environment. It is, therefore, prudent to account for the full range of structural variability in describing the geologic facies distribution by considering multiple TIs. Here, a Bayesian mixture model is presented for adaptive and efficient sampling of conditional facies from multiple uncertain TIs. The posterior distribution of facies is partitioned into individual conditional densities of the TIs and the corresponding mixture weights are estimated from the likelihood function for each TI. To implement the conditional sampling, a

---

* Part of this section is reprinted with permission from "A Bayesian mixture-modeling approach for flow-conditioned multiple-point statistical facies simulation from uncertain training images" by Khodabakhshi, M. and Jafarpour, B., 2013. Journal of WATER RESOURCES RESEARCH. 49, 1-15. Copyright 2012 by American Geophysical Union.

recently developed ensemble Kalman filter (EnKF)-based probability conditioning method is applied, whereby EnKF is used to invert the flow data and obtain a facies probability map (soft data) to guide conditional facies simulation from each TI. The suitability of the proposed Bayesian mixture-modeling approach is demonstrated using several numerical experiments in fluvial formations with uncertain orientation and structural connectivity.

4.2. Training Image Uncertainty

The main difficulties in understanding and modeling subsurface phenomena is related to inaccessibility and heterogeneity of geologic formations, together with the complex interactions between fluids and rocks over a wide range of temporal and spatial scales. Consequently, significant uncertainty is introduced into predictions of the related flow and transport processes, thereby complicating the development of subsurface hydrocarbon resources.

Multiple-point statistics (MPS) presents a grid-based pattern-imitating simulation method to model complex geological connectivity that is not amenable to variogram-based modeling techniques. A review of MPS sampling was presented in Subsection 1.3. However, a standing challenge in MPS-based model calibration is the uncertainty in the prior training image (TI). This issue becomes particularly important considering the strict pattern-imitating nature of MPS simulation that restricts the spatial variability of the resulting facies to the structural connectivity and encoded patterns in the given TI. Specifically, realization of facies maps from TIs with different structural connectivities

can exhibit distinctly different flow and transport prediction, which can be detrimental for development planning.

The main objective in this section is to develop an adaptive sampling strategy when multiple TIs are used to acknowledge the uncertainty in the geologic continuity model. A key question to address is how to identify and sample from relevant TIs in a list of candidate prior TIs. A Bayesian mixture-modeling algorithm is introduced for generating conditional facies realizations from multiple uncertain TIs. Data scarcity and low resolution, together with errors in geologic modeling and imperfect assumptions can leave significant uncertainty in interpretation of the existing patterns in a prior TI model.

One approach to deal with the uncertainty in describing the geologic continuity in a TI is to consider several TIs that capture the full range of geologic variability for a given formation. These TIs could be obtained based on different plausible geological scenarios, for example, from independent interpretations by different geologist or by stochastic treatment of parameters in a geologic modeling study that is used to identify possible connectivity patterns in the formation.

One way to apply the EnKF-based PCM to conditional facies simulation under several TIs is to randomly draw an ensemble of facies realizations from different TIs to account for the full range of variability in geologic continuity. Jafarpour and McLaughlin (2009b) applied the regular EnKF with an ensemble derived from fluvial channel TIs with different channel widths and showed that an ensemble composed of a mixture samples from different TIs may still be able to retrieve the overall continuity structure from flow data. A similar experiment can be set up using the PCM approach.

In Subsection 2.4.4.2, the effect of using an incorrect prior TI on the performance of regular EnKF and the PCM implementation was shown. In the next subsection, it is shown that uniform sampling from multiple TIs with different structural connectivity does not provide an effective solution to account for TI uncertainty. As an effective alternative method, a Bayesian mixture-model formulation is introduced for adaptive sampling from multiple prior TIs.

4.2.1. Example 4.1: Uniform Sampling from Multiple TIs

In the first example, numerical experiment in two-dimensional two-phase (oil-water) system is used. The model includes a nine-spot well configuration with one water injection well in the center and eight symmetrically located oil producers on reservoir boundaries. The initial oil saturation is 0.90 everywhere in the reservoir, whereas the initial pressure is $2000psi$ everywhere. A total of 0.7 pore volume of water is injected during 72 months of simulation. The production ports operate with a constant pressure of $1950psi$ for wells inside the high-permeability (channel) facies in the reference model and $1000psi$ for wells completed in the non-channel facies. Under these conditions, the injection pressure and water and oil production rates were measured and used as calibration data.

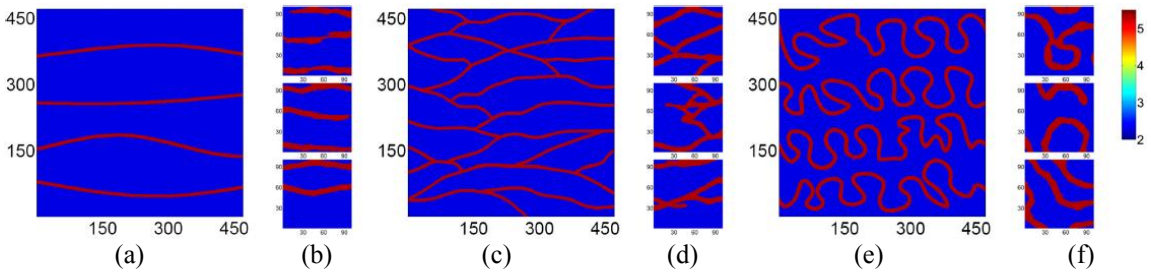


Fig. 4.1 – Three TIs with different structural connectivities and three corresponding unconditional MPS realizations from them: (a) the TI with straight channels and (b) corresponding realizations; (c) the TI with intersecting straight channels and (d) corresponding MPS realizations; and (e) the TI with meandering channels and (f) corresponding MPS realizations.

The measurements were obtained every 6 months by running the forward simulation with a specified reference permeability field. The channel and background facies are assigned permeability values of $200mD$ and $10mD$, respectively. The TIs for all the experiments in this section are shown in **Fig. 4.1**. **Fig. 4.2(a,b)** shows the reference model and well configurations used in the example 4.1.

One approach to incorporate the uncertainty in the prior TI in the above model calibration problem is to combine multiple TIs. That is, one could include more diverse features in one TI to capture a wide range of variability. This approach may require a very large TI if the stationarity assumption is to be preserved. In addition, in some geologic formations it may be hard to justify combining several alternative geologic patterns into a single TI. Therefore, except for investigating small variabilities in the conceptual model, combining multiple TIs with different structures into one does not provide a rigorous solution.

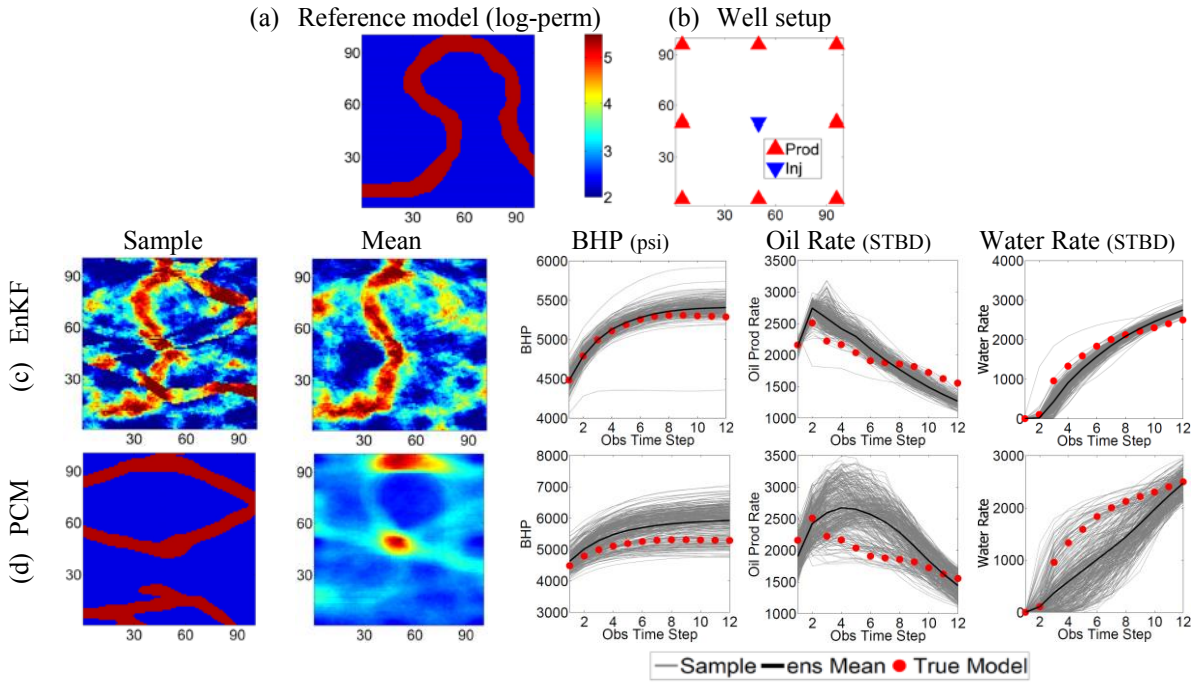


Fig. 4.2 – Facies estimation results for standard EnKF and PCM facies model calibration in Example 4.1: (a) true log-permeability field and (b) well configuration; The results for (c) EnKF and (d) PCM when an equal number of samples are used from each TI. One final facies sample (first column), the mean of final facies ensemble (second column), and the production forecast with final replicates (third to fifth column) are shown where the red circles show the observed response of the reference model; the forecasts with each log-permeability realizations are shown with thin gray lines, whereas the mean forecast is displayed with thick black lines.

As an alternative approach, to represent TI uncertainty it is possible to sample facies realizations from multiple TIs. In this case, it is critical to design an effective sampling strategy so that the number of samples drawn from each TI is proportional to the overall flow prediction performance of the samples taken from it which is achieved through an adaptive sampling mechanism, described in the next section. To show the

importance of an adaptive sampling strategy, three TIs (Figs. 4.1(a), 4.1(c), and 4.1(e)) are used and the EnKF and PCM calibration are applied by taking equal number of facies samples (100) from each TI. Note that the TI in Fig. 4.1(e) is the same TI from which the reference model is simulated.

Table 4.1 – RMSE and Spread for production forecast in Example 4.1.

	BHP		Oil prod rate		Water prod rate	
	RMSE	Spread	RMSE	Spread	RMSE	Spread
Initial	0.124	0.0794	0.3294	0.1797	0.7225	0.3673
EnKF	0.0209	0.0178	0.1052	0.055	0.1365	0.07
PCM	0.1158	0.0776	0.3207	0.1811	0.6747	0.3754

Fig. 4.2(c) has the results for the standard EnKF when 100 samples from each TI are used for model calibration. Although the estimated results are not discrete random fields, they contain important information about the location of channel facies. This observation is consistent with the results reported by Jafarpour and McLaughlin (2009b). Note that even though the updated log-permeability distribution does not share the discrete nature of the reference model, it contains important information about facies distribution, which is taken advantage of in the PCM approach (see Section 2). In addition, as shown in **Table 4.1**, the ensemble Spread and RMSE are both reduced after data assimilation despite the poor geologic reconstruction of the reference model.

Fig. 4.2(d) illustrates the calibration results using PCM with the same three TIs and by taking 100 samples from each TI. In this case, the majority of final facies

samples (2/3 of all samples) are taken from the incorrect TI. As expected, the forecasts with final replicates from three TIs with different structures exhibit wide variability. In this case, the forecast ensemble envelopes the observed data. The log-permeability ensemble mean also shows a slight trace of the meandering trend in the field. Clearly, when the three TIs have very different structures, only one (if any) of the TIs can be relevant; therefore, the uniform sampling strategy employed in this example becomes ineffective.

A Bayesian mixture model is combined with the probability conditioning method (PCM) for adaptive conditional sampling from multiple TIs. A detail review of PCM was presented in Section 2. The adaptive conditional sampling is accomplished by initially generating unconditional facies realizations from multiple TIs using an initially equal weight for each. The TI weights are then updated based on their predictive performance (likelihood function). For conditional sampling, the dynamic flow data is converted into a facies probability map using the PCM. The generated probability map is incorporated as input into MPS simulation to draw new conditional facies samples from each TI according to the weights assigned to each TI based on their likelihood to match the observed data. This leads to an adaptive facies sampling technique where fewer (more) realizations are generated from the TIs with inconsistent (consistent) geologic continuity. In the next section, an adaptive sampling strategy is proposed to address this issue.

4.3. Bayesian Mixture Modeling (BMM)

A Bayesian mixture-model formulation is presented for estimating TI importance weights, which will subsequently be used for adaptive sampling from the TIs. To this end, the mixture weights $p(T_j|y_{1:n})$ are considered such that $\sum_{j=1}^{J} p(T_j|y_{1:n}) = 1$ and $0 \leq p(T_j|y_{1:n}) \leq 1$. To compute the desired posterior density $p(m|y_{1:n})$, the individual posterior densities $p(m|y_{1:n}, T_j)$ for each TI T_j and the corresponding mixture weights $p(T_j|y_{1:n})$ must be calculated. Assuming that the facies model m is a mixture model of J different density functions (TIs here), the posterior density $p(m|y_{1:n})$ is

$$p(m|y_{1:n}) = \sum_{j=1}^{J} p(m|y_{1:n}, T_j)p(T_j|y_{1:n}) \tag{4.1}$$

In this ensemble formulation, the mixture weights determine the sampling weights for each TI and refer to the relative number of realizations taken from each TI. The individual posterior densities for a given TI can be expressed as

$$p(m|y_{1:k}, T_j) = p(m|y_{1:k-1}, y_k, T_j) = \frac{p(y_k|m, y_{1:k-1})p(m|y_{1:k-1}, T_j)}{p(y_k|y_{1:k-1}, T_j)} \tag{4.2}$$

where the equality $p(y_k|m, y_{1:k-1}, T_j) = p(y_k|m, y_{1:k-1})$ is used. The first and second terms of the numerator in equation (4.2) can be computed from the observation model and the posterior at the previous time step, respectively. The denominator in equation (4.2) is a normalization constant and is not trivial to calculate. In this ensemble framework, I am interested in drawing samples from the above posterior density function for each given T_j according to the mixture weights. The mixture weights $p(T_j|y_{1:n})$ for each TI can be computed by invoking the Bayes rule as follow:

$$p(T_j|y_{1:n}) = \frac{p(y_{1:n}|T_j)p(T_j)}{p(y_{1:n})} \tag{4.3}$$

where $p(y_{1:n})$ is independent of T_j and can be treated as constant, and $p(y_{1:n}|T_j)$ is the likelihood of observing the data under TI T_j, which is hard to compute unless simplifying approximations are applied. For example, under Gaussian posterior approximation, $p(y_{1:n}|T_j)$, may be computed in closed form using equation (4.2). Since the posterior is not Gaussian in these examples, an alternative approximation approach using Monte Carlo simulation is used.

For sequential conditioning in time, the likelihood can be computed at each time step using the sequential form

$$p(y_{1:n}|T_j) = \prod_{k=1}^{n} p(y_k|y_{1:k-1}, T_j) \tag{4.4}$$

which after substituting in equation (4.3) leads to

$$p(T_j|y_{1:n}) \propto p(T_j) \prod_{k=1}^{n} p(y_k|y_{1:k-1}, T_j) \tag{4.5}$$

To compute the mixture weights, the individual likelihood functions in equation (4.5) must be obtained. These likelihood densities can be calculated through integration over m, i.e.,

$$p(y_k|y_{1:k-1}, T_j) = \int_m p(y_k|m, y_{1:k-1}) p(m|y_{1:k-1}, T_j) dm \tag{4.5}$$

For most realistic problems, finding the exact solution of the integration in equation (4.5) is not feasible. However, Monte Carlo approximations can be used to find samples from

$p(y_k|y_{1:k-1}, T_j)$. This approach is followed in this section. The steps involved in sampling from the TIs using the PCM approach are outlined in the next subsection.

4.4. Implementation of BMM with Multiple TIs as Prior

The implementation begins by assigning equal weight $p(T_j) = 1/J$ to each component of the mixture model, i.e., TI (initial weight can be different if existing knowledge suggests so).

For time steps $t = 1{:}\, N_t$

A. Prediction step

- Draw N_{ens} realizations from the TIs based on the current TI weights (initially $1/J$) and probability map (initially homogeneous) and assign permeability values to each facies type.
- Solve multiphase flow equations for each realization of the permeability ensemble from the initial time step to predict the observed measurements at current time.

B. Update mixture weights and probability map

- Compute the sample weights for each realization of permeability according to the following proportionality:

$$w_t^{j,i} \propto p(y_t|T_j) \propto c \exp\left(-\frac{1}{2}\left(h_t(m^{j,i}) - y_t^{obs}\right)^T C_y^{-1}\left(h_t(m^{j,i}) - y_t^{obs}\right)\right) \qquad (4.6)$$

These weights will be used for adaptive sampling from the TIs for the next time step.

- Using the EnKF analysis equation, update the mean of permeability realizations by integrating the observations at the current time step.
- Use the mean permeability map in previous step to construct an updated probability map (following the procedure shown in **Fig 4.3.**).

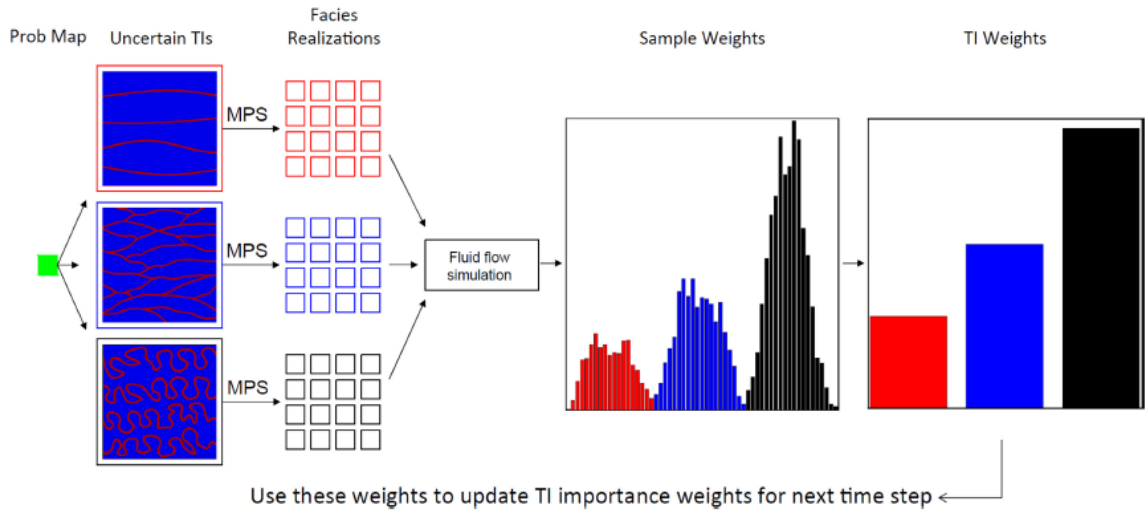


Fig. 4.3 – Schematic of steps involved in the proposed workflow for mixture-model formulation of conditional sampling from uncertain TIs.

Since the EnKF update equation requires parameter covariance and cross-covariance between parameters and predicted responses, the log-permeability ensemble is used to derive these statistics. The updated log-permeability ensemble mean is then used to obtain probabilistic information about spatial facies distribution. To do this, the probability of each facies is considered at a grid block to be linearly related to the

difference between the updated permeability and the value of permeability assigned to each facies. Hence, the update equation is only used to condition the permeability mean on flow data and the updated permeability mean is used to infer a probabilistic description for the facies distribution (a probability map). The updated probability map is used to select an updated ensemble of facies models to perform the sequence of EnKF forecast and update for the next time step.

Note that, following each update step, a new probability map is generated and used to draw an updated ensemble of conditional facies models from the available TIs. For parameter estimation, the states (pressures and saturations) are derived by rerunning the forward simulations from the initial time step using the updated facies models. This eliminates the loss of conservation principles (at the cost of a computational overhead) that is widely known for joint state and parameter estimation with EnKF. It is also worthwhile to note that the permeability updates are expected to be more accurate at the proximity of the wells and somewhat inconclusive away from the wells. The PCM approach is designed to combine the flow data with a prior TI model such that the former is used to deduce the local trends around the observation points (where flow data are more informative) and the latter to describe the facies connectivity away from observation points, where the data tend to be less conclusive (Section 2). When multiple TIs are used, the simulated facies samples tend to be similar around the well locations and become more dissimilar with increasing distance from the wells.

In the next section, the performance of the mixture-model PCM approach is demonstrated under TI uncertainty using three numerical experiments for adaptive

conditional sampling. Finally, note that EnKF is only used to update the facies probability map, whereas the likelihood function is used to update the TI consistency weights.

Table 4.2 – General Simulation/Assimilation Information

Parameter	*Special value/condition*
Simulation parameters	
Phases	Two-phase (oil/water)
Cell dimensions	$30ft \times 30ft \times 30ft$
Rock porosity	0.2 (uniform)
Initial oil saturation	0.9 (uniform)
Initial pressure	2000psi
Injection well constraints	Water flow rate
Production well constraints	Bottom hole pressure
Facies type	Fluvial formation
Geostatistical simulation	Snesim
Assimilation information	
Observation at the injection wells	Bottom hole pressure
Observation at production wells	Oil and water flow rate

4.5. Application

In this section, three examples are presented to investigate the uncertainty in structure and direction of channel facies in a fluvial formation. In the first two-dimensional example, the structural connectivity of the TI is uncertainly not known, whereas, in the second two-dimensional example, channel orientation for a given structural connectivity is considered unknown. In the third example, 5 layers of SPE10 benchmark are used as true synthetic model and the structural connectivity is uncertainly unknown. An adaptive

sampling strategy is used to identify the correct structural connectivity in Examples 4.2. and 4.4. and the consistent channel direction in Example 4.3. The MPS simulations in the following examples are carried out using the snesim implementation in the Stanford Geological Modeling Software (Remy et al.; 2009). The general simulation and data integration parameters used for all the experiments are summarized in **Table 4.2**.

4.5.1. Example 4.2: Channel Structure Uncertainty

Example 4.1 is repeated with uncertain structural connectivity. Three TIs (Figs. 4.1(a), 4.1(c), and 4.1(e)) are considered to reflect the uncertainty in the prior connectivity models. In the proposed mixture-model method, an ensemble with $N = 300$ realizations is used. Similar to the examples in Section 2, a simple linear mapping is applied to convert the mean log-permeability to a facies probability map (Fig. 2.4a). The lower and upper bounds C_{min} and C_{max} determine the level of confidence that is placed in the flow data (i.e., via the probability map). $C_{min} = 0.1$ and $C_{max} = 0.9$ are chosen for the examples that follow.

The results of estimating the meandering channel facies using the mixture-model PCM approach are shown in **Fig. 4.4**. Figs 4.4(a) and 4.4(b) show the reference facies map and the field setup. Fig 4.4(c) illustrates the changes in TI consistency weights with each EnKF analysis step. The weight associated with the correct TI (with meandering channels) is shown with a solid blue line. Within about five EnKF updates, the consistent TI is correctly estimated, and almost all facies realizations are taken from the consistent meandering TI. Comparing the results from the mixture model approach (Fig. 4.4(f)) with those from the EnKF and PCM updates in Fig. 4.2 reveals that the

probability map estimated by the mixture-model approach better represents the meandering feature in the true facies map. The level of variability in the simulated facies ensemble can be adjusted by the bounds specified for the probability map and/or by changing the τ value in equation (1.3).

Without adaptive sampling, the standard PCM approach uses a large number of samples from the two inconsistent TIs. As a result, the sample covariance that is used in the EnKF update becomes inaccurate and degrades the quality of the update (Fig. 4.2). However, by updating a weight that represents the TI consistency in reproducing the flow data, the proposed approach can continuously improve the EnKF analysis, since the updated covariance matrix after adaptively conditional resampling from the TIs better represents the spatial connectivity and correlation in the field. It only takes about four updates in this example to identify the correct TI.

However, even with after identifying the correct TI, some uncertainty, mainly about channel boundaries, still remains in the calibrated ensemble of facies models (see Fig. 4.4, last row). This uncertainty may be attributed in part to the scattered nature of the data and the weak sensitivity of available flow and pressure data to the exact location of the channel boundaries.

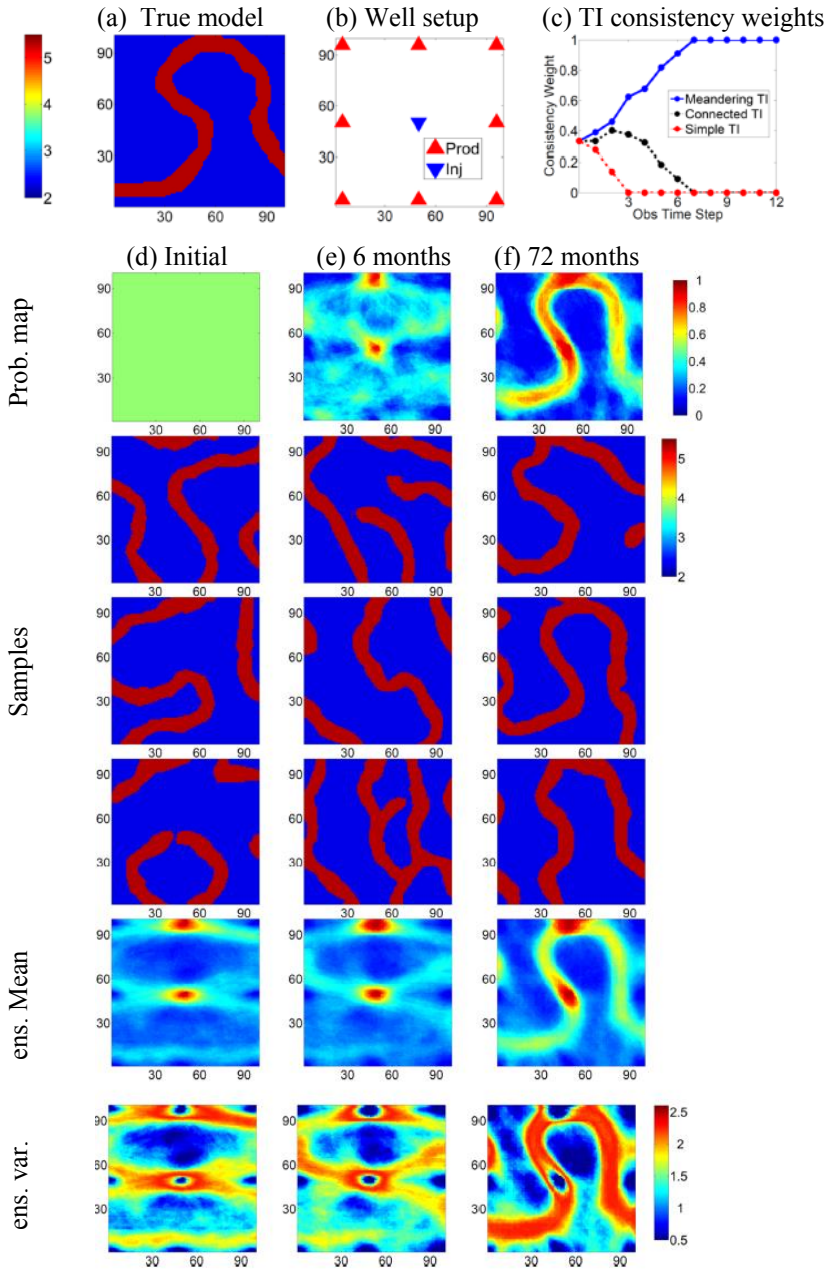


Fig. 4.4 – Facies estimation results for mixture-model-based PCM approach in Example 4.2: (a) true log-permeability field, (b) its corresponding well configuration, and (c) evolution of the TI consistency weights throughout data integration steps. (d)–(f) The probability map (first row), three sample log permeabilities (second to fourth rows), and the mean and variance of ensemble log permeability (fifth and sixth rows, respectively) are shown for the initial time, after 6 months, and 72 months of data assimilation, respectively.

The production forecast performance is measured using normalized ensemble-based root-mean-square error (RMSE) and normalized ensemble spread defined as follows:

$$RMSE(X) = \frac{1}{n_{obsT} \times n_{well}} \sum_{t=1}^{n_{obsT}} \sum_{w=1}^{n_{well}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(x_i^{t,w} - x_{true}^{t,w}\right)^2} \tag{4.7}$$

$$Spread(X) = \frac{1}{n_{obsT} \times n_{well}} \sum_{t=1}^{n_{obsT}} \sum_{w=1}^{n_{well}} \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(x_i^{t,w} - \bar{x}_{ens}^{t,w}\right)^2} \tag{4.8}$$

where n_{obsT} is the total number of observation times, n_{well} is the total number of wells, and N is the ensemble size. The spread and RMSE values become very similar when the estimated forecasts become unbiased (forecast mean approaches the true forecast).

Fig. 4.5 summarizes the results from predicting the reservoir dynamic response. The results are also presented in terms of spread and RMSE performance measure in **Table 4.3**. It shows that before calibration the flow response of the initial ensemble deviates from the observed measurements, whereas the forecast with the calibrated ensemble better follows the observed trend in the data. Compared with Fig. 4.2(c), the proposed approach does not underestimate the ensemble spread. This is attributed mainly to the resampling step after each data integration step to generate new facies models.

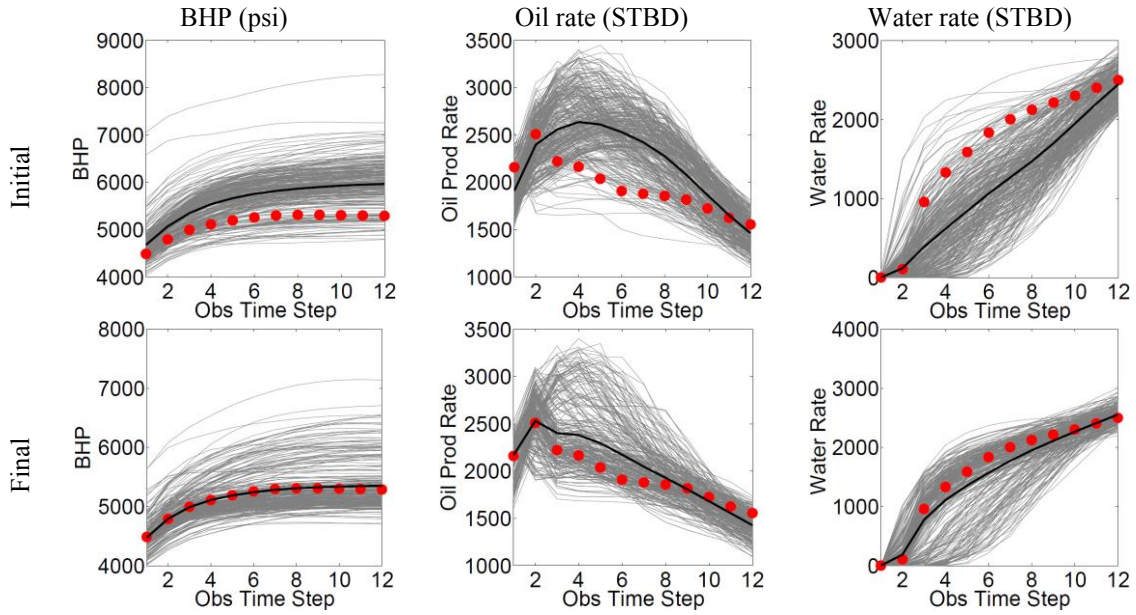


Fig. 4.5 – (top) Initial and (bottom) final production forecasts in Example 1. The first column (left) shows the injection bottom hole pressure, and the second and third columns display field oil and water production rates, respectively. Observations are displayed using red circles, while the forecasts for individual realizations and their mean are shown with thin gray and thick black lines, respectively.

It is also important to recognize that, similar to other probabilistic sampling techniques, PCM may generate conditional facies realizations that do not reproduce the observed flow response (Section 2). From Table 4.3, it can be confirmed that, after data assimilation, the RMSE of the flow predictions is significantly improved, whereas the spread is also slightly reduced. Interestingly, the reduction in spread is not as significant. A comparison between the RMSE and spread after data assimilation shows that the two measures become very similar, implying that the forecast estimates after calibration are less biased. Assigning higher τ values in the MPS simulation or changing the bounds

(C_{min} and C_{max}) in the probability map to increase the similarity between the generated samples are mechanism that can be used to improve the sampling results and reduce the ensemble spread.

Table 4.3 – RMSE and Spread for production forecast in Example 4.2.

	BHP		Oil prod rate		Water prod rate	
	RMSE	Spread	RMSE	Spread	RMSE	Spread
Initial	0.124	0.0794	0.3294	0.1797	0.7225	0.3673
Final	0.0935	0.0896	0.2172	0.1728	0.3925	0.3367

4.5.2. Example 4.3: Channel Orientation Uncertainty

The proposed method is applied in a separate example to identify channel orientation in a fluvial system. For this example, the channel facies structures are drawn from the TI shown in Fig. 4.1(c) but assume that channel direction is unknown. This was done by using the same TI and specifying a rotation angle for the simulated facies realizations in the snesim algorithm. The channel direction was allowed to vary in the range $[-90°\ \ 90°]$, which covers all possible channel orientations.

The reference model and well configuration for this test case are depicted in **Fig. 4.6(a)** and 4.6(b), respectively. The channel direction in the correct TI changes between $-20°$ and $20°$, based on a histogram obtained from 1000 samples. Based on the directional variability in the TI, five specific directions, $-72°$, $-36°$, $0°$, $36°$, and $72°$ are used to cover the entire range of channel directions. With this specification, the

problem is reduced to adaptively sampling from the TI using the five rotation angles mentioned above. The total simulation time is 36 months. The simulation and model calibration parameters are kept the same as in previous example.

The results of applying the mixture-model PCM to this problem are given in Fig. 4.6. Fig. 4.6(c) shows the evolution of the TI directional consistency weight after each data integration step. The weight for the correct (horizontal) direction consistently increases with each update. The red line in the third row of Fig. 4.6(c) indicates the main continuity direction for the reference model ($-6°$). After eight steps of assimilation, only the consistent TI has significant contribution. The final results (Fig 4.6(f)) show that the mixture-model PCM identifies the correct directionality and provides an estimate for the probability map that is consistent with the features in the reference model. The fourth and fifth rows in Figs. 4.6(c)–4.6(e) show two samples (out of 300) facies at different time steps. The improvement in the directionality of these facies samples is quite evident. Some (about 15%) of the sample facies at the final step do not have the correct structure in the reference model. One explanation for this variability, beside the randomness introduced during snesim sampling, is the small number of observations in space (five locations). Note that since the samples are conditioned on hard data, the variance at the well location is zero. The final variance map shows that the main uncertainty is associated with channel edges and at locations far from the observation points.

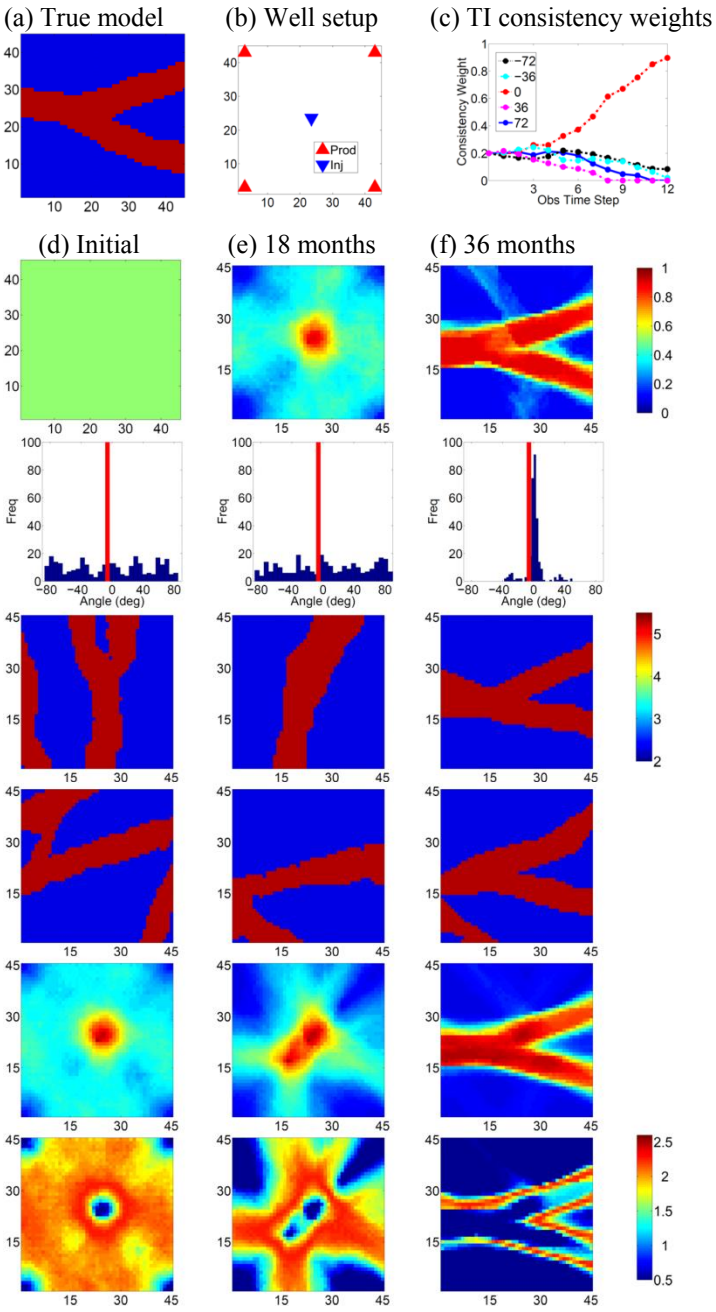


Fig. 4.6 – Facies estimation results for mixture-model-based PCM approach in Example 4.3: (a) reference log-permeability field; (b) its corresponding well configuration; and (c) evolution of TI consistency weights throughout data integration steps. (d)–(f), The probability map (first row), histogram of channel directions (second row) with the reference direction indicated with the red line, two sample log permeabilities (third and fourth rows), and the ensemble log-permeability mean and variance (fifth and sixth rows, respectively) are shown for the initial step, after 6, and 12 data assimilation steps, respectively.

In the above examples, the reference model belongs to one of the prior TIs. Since the TIs had different features in them, the correct TI is identified as a single consistent TI. In some cases, the reference features in the solution may be adequately captured with more than one of the prior TIs or, in a more pessimistic but possible case, with none of the TIs. In such cases, the TI consistency weights and the probability map may be used as useful pieces of information for designing a new TI based on preliminary calibration results. When neither of the TIs provides an adequate representation of reality, existence of a large bias in the forecast ensemble may hint at a fundamental problem in the data integration process and the prior model used. In a more problematic situation, the ill-posed nature of the problem may result in incorrect geologic models that explain the sparse data reasonably well. In such cases, additional data are required to diagnose possible errors in the solution. Ultimately, the quality of any inversion method is impacted by the validity of the prior model used. This becomes even more important when the solution is constrained to preserve the higher-order statistics of the prior.

4.5.3. Example 4.4: SPE10 Five-Layer Model

The mixture model PCM is applied with a two-phase three-dimensional model to investigate the efficiency of the proposed model in 3D system. Five layers of SPE10 benchmark are used as true synthetic model. The model includes a five-spot well configuration with one water injection well in the center and four symmetrically located oil producers on the corner of the reservoir. The initial oil saturation is 0.8, everywhere. The initial pressure at $12000ft$ (the datum of the reservoir) is $6000psi$. The injector and producers operate with a constant injection rate of $2500STBD$ and production pressure

of $4000psi$, respectively. Under these conditions, the injection pressure and water and oil production rates were measured annually months for 4 years and used as calibration data. The five layers of the true synthetic model are illustrated in **Fig. 4.7(a)**.

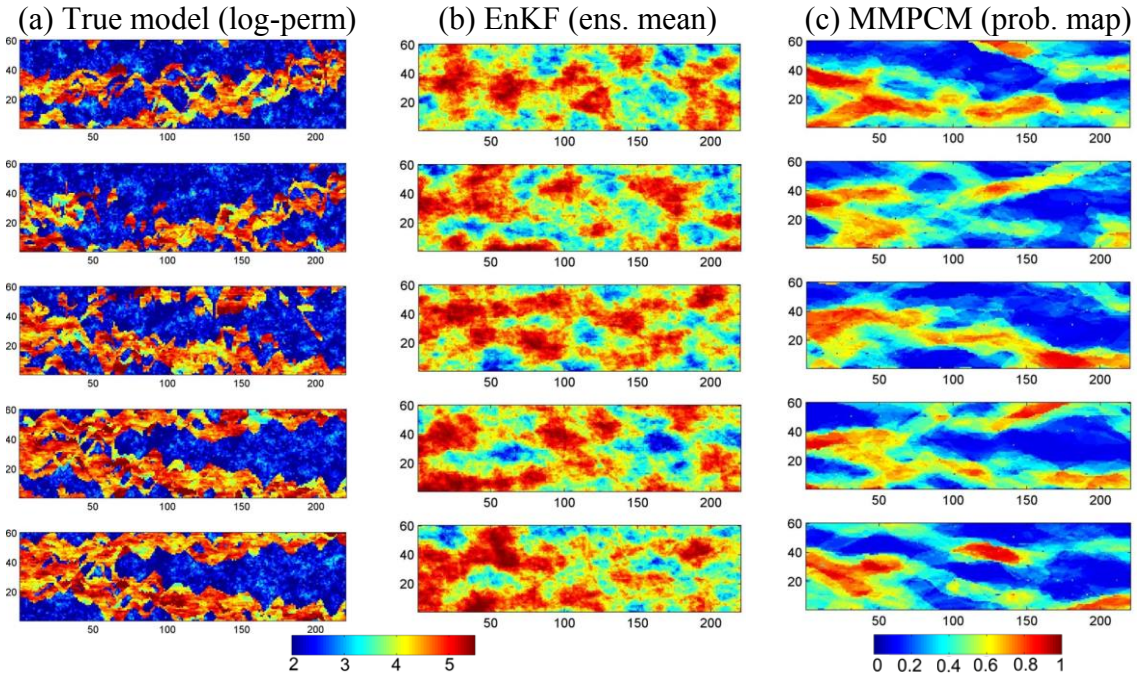


Fig. 4.7 – Facies estimation results for standard EnKF and the mixture model PCM approach in Example 4.4. (a) True log-permeability field (all five layers obtained from SPE10 benchmark), (b) final mean of log-permeability ensemble updated with standard EnKF, (c) final probability map of channelized facies updated using mixture model PCM.

Two TIs (Figs. 4.1(c) and 4.1(e)) where considered to reflect the uncertainty in the prior model. As illustrated in Fig. 4.7(a), the synthetic model has facies structure which is locally meandering (like Fig. 4.1(e)) and globally connected channels (like Fig.

4.1(c)). Fig. 4.7(b) has the results for the standard EnKF when 100 samples from each TI are used for model calibration. Similar to previous examples in this section and Section 2, the estimated results are not discrete random field. Fig. 4.7(a) illustrates the calibration results using mixture model PCM with the same two TIs and by taking 100 samples from each TI at the first step.

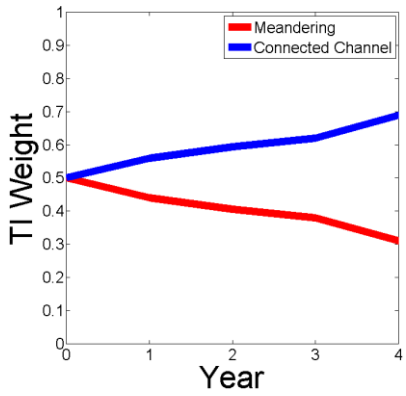


Fig. 4.8 – Evolution of TI consistency weights throughout data integration steps using mixture model PCM for Example 4.4.

Fig. 4.8 shows the evolution of the TI consistency weight after each data integration step. Even though neither of those TIs is perfectly consistent with the underlying features in true synthetic model, MMPCM provides higher consistency weight for TI with connected channel which globally more consistent with true model. To investigate the effect of incorporating geological information during the history matching process, the rate optimization is performed using the true model, the best

model updated using standard EnKF, and the best updated using mixture model PCM. Then, forecast for the next 6 years is obtained using true model and optimized production/injection rate controls obtained from each model.

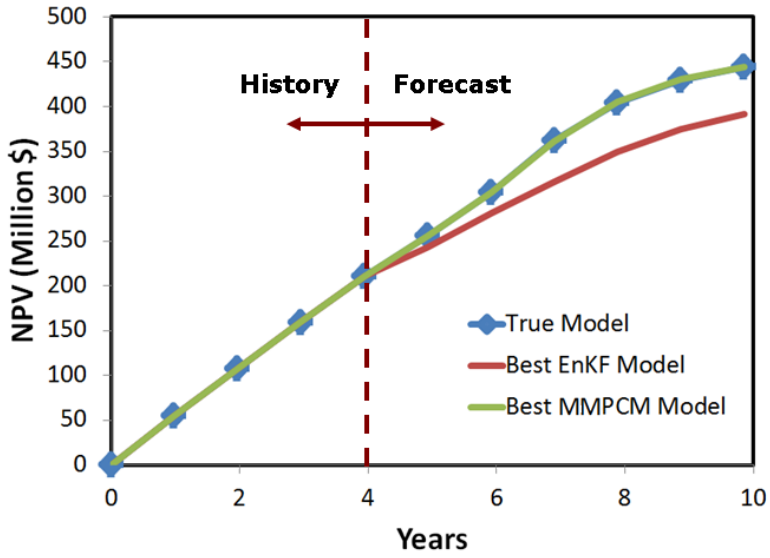


Fig. 4.9 – NPV in time using optimized rate control obtained using true synthetic model (blue line), best EnKF model (red line), and best MMPCM model (green line) for Example 4.4.

Fig. 4.9 illustrates the net present value (NPV) with time for these three models. As depicted in Fig 4.9, the NPV obtained from optimizing the rate using the best model updated by mixture model PCM has similar trend as the NPV obtained from true synthetic model. At the same time, optimizing rate controls using the best model updated by standard EnKF provides significantly smaller NPV in time. The ratio of $NPV_{MMPCM}/$

$NPV_{True\ Model}$ is equal to 0.98 at the end of forecast time (6 years), while $NPV_{EnKF}/NPV_{True\ Model}$ is equal to 0.88. In other words, using final replicates updated by MMPCM approach for rate optimization provides 10% more NPV in compare to using final reservoir models updated by EnKF for the rate control optimization.

Fig. 4.10 shows the final oil saturation for a selected layer (the third layer) at the end of 10 years of production for true synthetic model, best EnKF model, and best MMPCM model. The well controls are set to the optimized controls obtained for each individual model. As depicted in Fig. 4.10(top), there is an unswept area in the east of the true model which could be a good candidate for further field development. The final saturation map for the best MMPCM model shows the same unswept area while the best EnKF model fails to predict the correct oil saturation map.

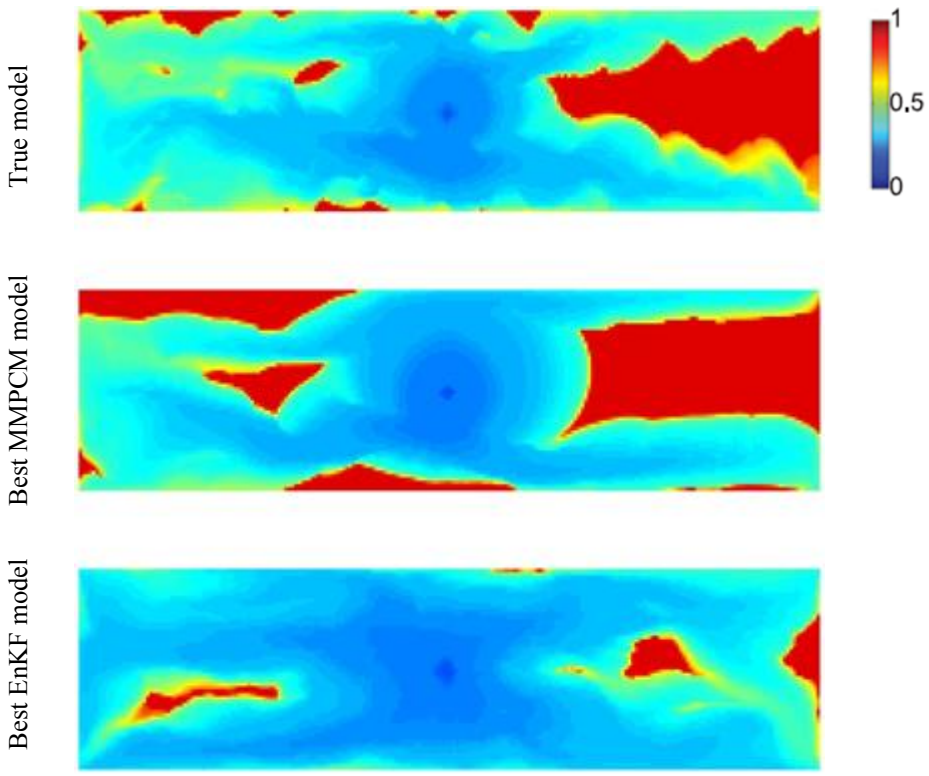


Fig. 4.10 – Final oil saturation map calculated using (top) true model, (middle) best MMPCM model, and (bottom) best EnKF model for Example 4.4. The rate controls are set to the estimated optimized rate controls for each individual model.

4.6 Conclusions

A Bayesian mixture model is presented for adaptively conditioning the simulation of geologic facies from multiple TIs to nonlinear pressure and flow measurements. The mixture model approach uses the PCM data integration algorithm that is described in Section 2 to adaptively draw conditional facies realizations from multiple TIs. The PCM is applied to convert dynamic pressure and flow data into facies probability map, which,

in turn, was used to guide facies simulation from prior TIs. At each sampling stage, the TI weights were estimated based on the likelihood function for the individual prior TIs and determined the number of samples drawn from each. After presenting the implementation details of the proposed approach, its performance is examined under several prior assumptions, including uncertainty in formation connectivity (2D and 3D model) and channel direction (orientation) for a given TI in fluvial systems. When channel direction was uncertain (Example 4.3), more update steps were needed for the algorithm to identify the correct directionality in compare to the number of steps necessary to detect the correct structure in the Example 4.2. This can be attributed to at least two main differences between these problems. First, in the example with structural uncertainty, the TIs had the same global directionality (horizontal), unlike in the example with unknown channel direction. Hence, the sample covariance calculated in the latter included structural correlations from all directions, resulting in a more pronounced degradation of covariance. The second major difference between the two problems is related to the location of the channel features in the reference model relative to the well configuration. In Example 4.3, the channel feature is less directly observable from the data as it is intersected by fewer wells. In all cases, however, the method was eventually able to identify, and accordingly sample from, the correct TI. In the last example of this section, the efficiency of the proposed approach is investigated for a two-phase three-dimensional problem under structural uncertainty. The optimized well controls estimated using the best MMPCM model and true synthetic model provide similar NPVs, while optimized control estimated using the best standard EnKF results in

10% less NPV. More importantly, the similar unswept area in the oil saturation maps obtained from the best MMPCM and true synthetic model offer a good candidate for infill drilling. This unswept area is missing in the saturation prediction using the best EnKF model.

Although the pattern-imitating nature of MPS simulation from prior TIs presents an opportunity to model more complex geologic phenomena, it also poses an important risk when the prior TI fails to represent the correct facies connectivity. Since the resulting facies models often dominate fluid displacement behavior, it is important to take into consideration the TI uncertainty in MPS simulation. Conditioning the MPS simulation results on nonlinear dynamic data remains an important topic in application of this method to modeling complex subsurface systems. The complexity of conditional simulation problem increases when the uncertainty in the TI model has to be acknowledged and incorporated. When prior knowledge about the structural connectivity is not adequate to overwhelmingly support the use of a single TI, it is prudent to consider a wider range of possible structural connectivity models (TIs) and rely on the dynamic flow data to distinguish between alternative TI candidates and adaptively sample from them based on their predictive performance (likelihood to reproduce the observed data).

Up to this point, the data integration is discussed using fine scale model. In the next section, a new workflow for integration of pressure data for estimating large-scale reservoir connectivity is proposed. According to the spatial resolution of the observations, the unstructured grid system is generated through Delaunay triangulation by using the location of the static pressure measurements as control points. The

important advantages of the proposed workflow for characterization of field-scale reservoir connectivity from pressure data include very fast connectivity estimation with a low-order model and effective parameterization to reduce the number of unknowns to a level commensurate with the available static pressure measurements. Successful application of the proposed approach to data from real fields illustrates its suitability and application to realistic reservoirs. The compatibility of the estimation results with the existing geological evidence verifies the performance of the proposed method.

5. MULTISCALE MULTI PHYSICS HISTORY MATCHING

5.1. Summary

Reliable early characterization of global reservoir connectivity is critical for improving field development plans. A main difficulty in identification of field-scale reservoir connectivity is the discrepancy between the low resolution of available field-scale pressure data and the high resolution of geologic models. In this section, a workflow is proposed for integration of pressure data for estimating large-scale reservoir connectivity. Since pressure variation represents a smooth function, an extremely low-resolution (coarse scale) grid system is adopted for reservoir simulation. The grid system is generated through Delaunay triangulation by using the location of the static pressure measurements as control points and the unstructured grid blocks are distributed according to the spatial resolution of the observations. Using flow-based upscaling, the initial coarse-scale static simulation model is created from the fine-scale geological model. Then the ensemble Kalman filter (EnKF) is used to automatically adjust the global parameters such as aquifer strength, global continuity/discontinuity of reservoir properties, and fault transmissibilities to match the static pressure. The important advantages of the proposed workflow for characterization of field-scale reservoir connectivity from pressure data include very fast connectivity estimation with a low-order model and effective parameterization to reduce the number of unknowns to a level commensurate with the available static pressure measurements. This framework is successfully applied for data from real fields to illustrate its suitability and application to

realistic reservoirs. The compatibility of the estimation results with the existing geological evidence verifies the performance of the proposed method.

5.2 Introduction and Significance

Interwell connectivity evaluation provides valuable information about the reservoir characteristics, which is useful in infill drilling and reservoir management optimization. For instance, a low estimate for interwell connectivity could be an indication of a barrier to flow like a sealing fault. On the other hand, large interwell connectivity values may be used as an evidence for fluvial channels or open fractures. In waterflooding, the observed injection and production rates can be used to estimate the connectivity between injectors and producers. A review of previous interwell connectivity estimation approaches was presented in Subsection 1.2.

In general, the resolution of the available data is not on par with the geologic model. In addition, the information content of each data type may correspond to a different scale or resolution of the model parameters. For example, pressure variation is a much smoother function of location than fluid front. To capture the fluid front, a fine-resolution grid system must be generated. In contrast, pressure distribution could be estimated using coarser grid blocks. The static pressure is usually not affected by local properties around the well. Hence, the pressure obtained from shut-in and/or new drilled wells provides valuable and useful information about large-scale reservoir continuity/discontinuity and the strength of external sources such as an aquifer. Analysis

of the pressure data can be used to characterize field-scale reservoir connectivity and compartmentalization for improved field development and management.

In this section, a novel approach is proposed for estimating field-scale reservoir connectivity by reconciling low-resolution field pressure data and high-resolution geologic models. The smooth pressure variation in the field lends itself to low-resolution (coarse-scale) grid representations and fast reservoir simulation for pressure prediction. To reconcile data and model resolutions, Delaunay triangulation is used to generate the grid system by taking the location of the static pressure measurements as control points and distributing the unstructured grid blocks according to the spatial resolution of observations. An initial static simulation model is generated by upscaling the high-resolution geological model to this coarse-scale unstructured grid system. The upscaled description in the unstructured grid system is used with an efficient data integration framework to delineate the reservoir connectivity. The ensemble Kalman filter is applied to match the static pressure by automatically adjusting the global parameters such as aquifer strength, global continuity/discontinuity of reservoir properties, and fault transmissibilities. The important advantages of the proposed workflow for characterization of field-scale reservoir connectivity from pressure data include very fast connectivity estimation with a low-order model, effective parameterization to reduce the number of unknowns to a level commensurate with the available static pressure measurements, and a multiscale implementation based on the resolution of (global versus local) production data.

To verify the performance of the proposed method, the proposed approach is applied for field data from naturally fractured reservoirs. The reservoirs are below the saturation pressure, so the reservoir fluid is considered compressible. In addition, the reservoirs are produced primarily under drainage with partial support from an adjacent aquifer. I demonstrate the compatibility of the estimation results with the existing geological evidences. This section begins with a presentation of the proposed method followed by a few examples to illustrate the application of the method and discuss its main properties.

5.3. Methodology and implementation

The global reservoir connectivity estimation from static pressure data is implemented in two steps. First, an unstructured coarse-scale grid system is generated to describe the connectivity in the fine-scale prior geological model at a scale consistent with the resolution of the available data. In the second step, the EnKF algorithm is used to estimate, from static pressure data, the parameters of the coarse-scale model reservoir that represent the flow connectivity in the reservoir.

Table 5.1 – General parameters used for the synthetic model

Parameter	Value/Condition	Parameter	Value/Condition
Phases	Oil and water	Simulation time	48 months
Fine grid system	$90 \times 90 \times 1$	Coarse grid system	5
Cell size (fine)	$30 \times 30 \times 30 ft^3$	Rock porosity	0.25
Initial oil saturation	0.8	Initial pressure	$6500 psi$
Injector constrain	Water flow rate	Producer constrain	Total fluid rate
# producers	4	# injector	1
# shut-in pressure	14		

5.3.1. Static Pressure Forecast

This section begins by describing the coarse-scale grid generation. A simple 2D test case with a 5-spot well pattern (one injection in the center and four producers in the corners of the model) is used to illustrate the coarse grid generation concept. This simple model simulates a waterflooding example with a two-phase immiscible (oil/water) system (**Fig. 5.1(a)**). The static pressures are estimated using the coarse-scale model and compared with those obtained from the fine-scale simulation. The initial oil saturation is uniformly distributed in the reservoir with a value of 0.80. The initial reservoir pressure is set at 6500 psi in all cells. Each producer operates under total rate control with the target rate of 100 STBD.

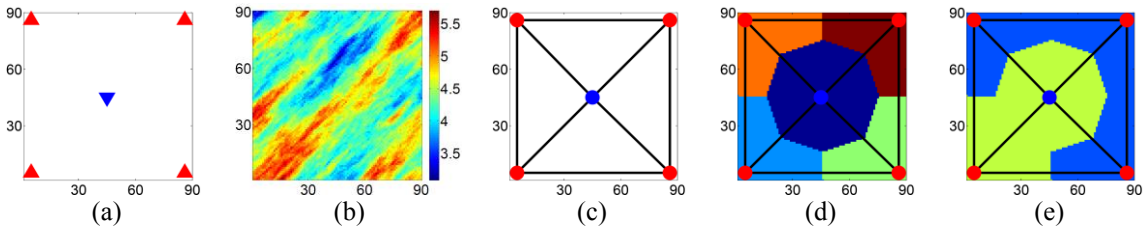


Fig. 5.1 – (a) Five-spot well configuration with one injector in the middle and four producers at the corners, (b) true log-permeability distribution, (c) Delaunay triangulation of the given well configuration in the first example, (d) unstructured grid blocks generated by connecting the center of the triangles to the middle point of each edge. Each color represents one grid block. (e) Active (green) and inactive (blue) fine-scale grid cells use to calculate transmissibility between coarse scale grids.

After each 90 days, one of the producers is randomly selected and shut-in for 30 days. The shut-in pressure measurement is obtained by measuring the bottomhole pressure (BHP) at the end of the shut-in period. The 120-day cycle is repeated 12 times. Also, two shut-in periods are set for the injector during the shut-in period of producers. Similarly, the BHP of the injector is obtained at the end of the shut-in period. The injector is constrained to inject the total production target rate of the producers. The main continuity in the permeability field is in the southwest-northeast direction (Fig 5.1(b)). The range of permeability values is 22md to 299md. **Table 5.1** summarizes the simulation parameters.

5.3.1.1. Coarse Scale Grid Generation

Reservoir connectivity is estimated from field pressure data. Since pressure varies smoothly in space, a coarse grid system is adopted for reservoir simulation. The grid system is proposed to reconcile the data and model resolutions. The Delaunay triangulation approach is adopted to generate the coarse grid. The locations of the static pressure measurements are used as control points and the unstructured grid blocks are distributed according to the spatial observation resolution. The Delaunay triangulation for a set of control points is a triangulation such that every circumcircle of a triangle is an empty set of the control points (Okabe et al. 1992). Since this triangulation approach maximize the minimum angles of all triangles, the generated triangles tend to avoid skinny triangles. Also, the method guarantees that the triangles are not overlapping, and every edge is shared by at most two triangles. However, in these cases a Delaunay triangulation is not guaranteed to exist or be unique. Delaunay triangulation is usually

adopted to generate meshes for space-discretized solvers such as the finite-element and the finite-volume methods. In this approach, the locations of static pressure data are the set of control points. The Delaunay triangulation of this well configuration is illustrated in Fig 5.1(c). Unstructured grids are generated by connecting the center of each triangle to the middle point of each edge. Fig. 5.1(d) depicts the unstructured grid blocks associated with this example. Each color represents one grid block. Using the proposed methodology, the grid blocks are distributed according to the spatial resolution of the available measurements.

The volumetric average of porosity of the fine-scale grids located inside each coarse grid block is used to upscale porosity from the fine-scale grid to the coarse-scale grid system. Using the volumetric average guarantees that the coarse-scale and fine-scale grid systems have the same pore volume.

However, permeability upscaling is more involved than porosity upscaling and it has been an active area of research (Wen and Gomez-Hernandez, 1996; Renard and de Marsily, 1997; Farmer, 2002; Durlofsky, 2005; Gerritsen and Durlofsky, 2005). Methods may be categorized as permeability or transmissibility upscaling approaches. Permeability upscaling methods provide equivalent permeability tensors for each coarse-scale grid block. Therefore, the interface transmissibilities between adjacent grid blocks are internally calculated by numerical simulators using the equivalent permeabilities of coarse grid blocks. In contrast, transmissibility upscaling techniques are involved in determination of coarse-scale transmissibility at the interface using upscaling

procedures. Transmissibility upscaling has provided more accurate coarse-scale models than permeability upscaling (Chen et al., 2003; Romeu and Noetinger, 1995)

In the proposed approach, flow-based transmissibility upscaling is used to transfer flow properties of the fine-scale to a coarse-scale reservoir model. To calculate the transmissibility between each pair of adjacent coarse-scale grid blocks, a flow-based upscaling approach is adopted. The entire fine-scale grid blocks except the ones located inside the two adjacent coarse grid blocks are inactivated. For example, to calculate the transmissibility between the coarse grid blocks associated with the injector and the bottom-left-corner producer, the fine grid blocks within these two coarse grid blocks remain active and the rest of the fine-scale grid blocks are inactivated; the corresponding active cells of the fine-scale grid systems are shown in the green regions in Fig. 5.1(e). The injector and producer are then set to a specified rate constraint. The volumetric average pressure inside each coarse scale grid is obtained by running forward simulations with the fine-scale system. The transmissibility between these adjacent grid blocks can be obtained from Darcy's law from the average pressure and flow rate. By repeating the same process, the transmissibility between each adjacent pair of coarse-scale grid blocks is obtained. This procedure yields the static properties for the coarse-scale grid system.

5.3.1.2. Forward Simulation Using both Fine- and Coarse-Scale Grid Systems

Forward simulation is performed using the fine-scale grid system described in Section 5.3.1.1 as the base case. The BHP values obtained for the injector and each producer are depicted as red dots in **Fig. 5.2(a-e)**. The shut-in period could be detected as the big

jump up (or down) in producer/injector BHP. The small jumps in the injector BHP are related to the change in the injection rate. The red dots in Fig. 5.2(f) illustrate all the shut-in pressures in time obtained from forward simulation of the fine-scale model.

To perform the forward simulation for the coarse-scale model, the ECLIPSE commercial reservoir simulator provided by Schlumberger is used. Since the transmissibilities between each pair of the coarse grid blocks is known, a one dimensional model is defined and each pair is connected using non-neighbor connection (NNC) capability in ECLIPSE. The injector and producers controls are set in the coarse model as the total fluid rate constraints. The forecast BHP values for injector and producers are depicted with the blue line in Fig 5.2(a-e). Even though the flow BHP of the two systems behaves differently, the values of static pressure obtained from the two models are similar. All the static pressures estimated using the coarse-scale model are illustrated in Fig. 5.2(f) (blue line). The RMSE between the base case (fine-scale model) and the coarse-scale static pressure estimation is 0.135 STBD. This result shows that the coarse-scale model can estimate a static pressure trend similar to the fine model using a small number of grid blocks. In this example, the upscaling factor is 1620 (90×90/5).

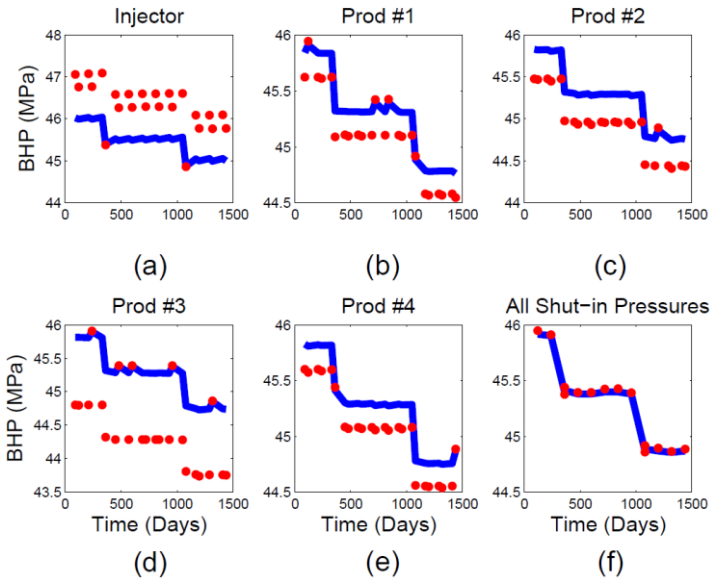


Fig. 5.2 – (a-e) Bottomhole pressure (BHP) of fine-scale (red dot) and coarse-scale model (blue line) of injector (a) and all producers (b-e) with time. (f) All static pressure with time for fine-scale (red dot) and coarse-scale model (blue line).

EnKF is used as a data assimilation approach. (The review of EnKF was presented in Subsection 1.4.1.) A major consideration in practical application of EnKF to large-scale problems with computationally demanding forward models is ensemble size. In general, a large number of realizations are needed to represent the variability in the forecast states and to approximate the required sample statistics to perform the EnKF analysis. However, for computationally demanding forward models, the ensemble size cannot be large, which often results in sampling errors and introduction of non-physical covariances. Since the proposed workflow uses very coarse grid blocks with fast forward-simulation runs, EnKF can be combined with the algorithm rather efficiently.

In the next section, the reservoir continuity for two field cases are resolved by combining the described coarsening approach with the EnKF data assimilation algorithm. The assimilated data is the static pressure measurements.

5.4. Field-Scale Applications

To motivate the proposed approach, two naturally fractured reservoirs are used and the proposed approach is applied to infer the global reservoir continuity/discontinuity, fracture density distribution, and aquifer strength. First, using the method described in the previous section, I generate coarse-scale simulation models. The EnKF analysis is used to automatically adjust the global parameters such as aquifer strength, global continuity and/or discontinuity of reservoir properties, and fault transmissibilities to match the static pressure. Finally, the performance of the method is examined by comparing the estimation results with the existing geological evidence.

5.4.1. Field Descriptions of Reservoirs A and B

In this section, a dataset obtained from a naturally fractured porcelanite oil reservoir is used. Reservoir A has produced for over three decades. This reservoir consists of multiple units over a vertical interval of approximately 10,000 ft and contains layers of shale. The production history, average reservoir pressure in time, and well counts are depicted in **Fig. 5.3**. As depicted in Fig. 5.3(a), gas production rate (and also production gas oil ratio) has a jump in time when the reservoir pressure went below some critical pressure.

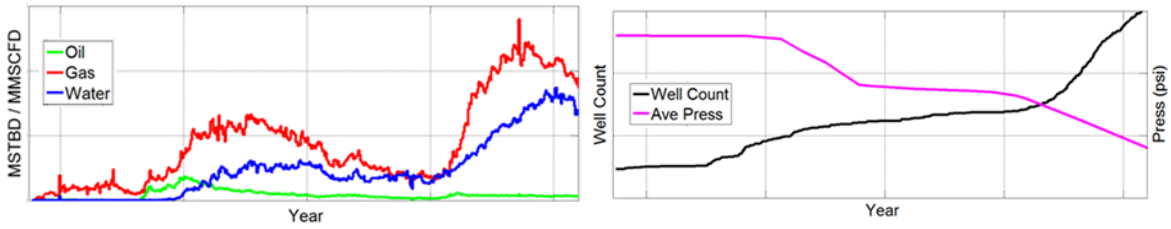


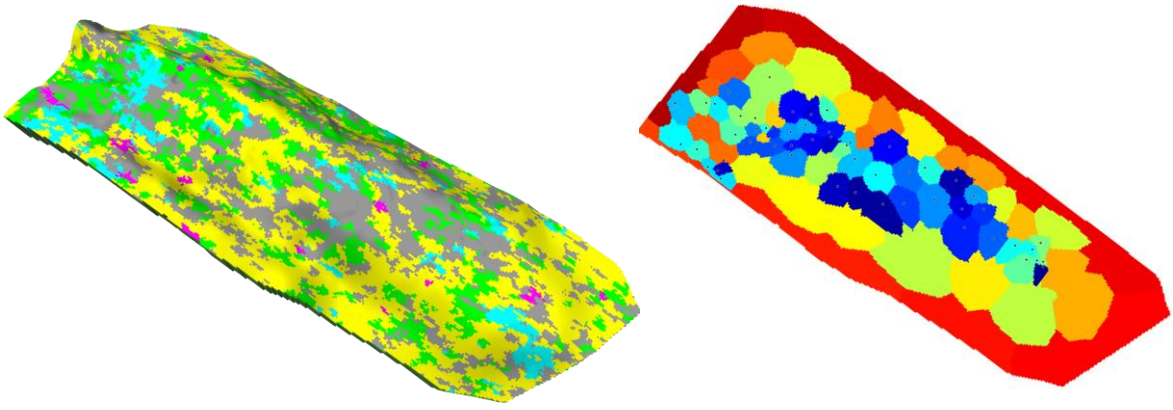
Fig. 5.3 – (a) Field production history and (b) average reservoir pressure and well counts with time for the reservoir used for Multiscale Multi-Resolution history matching (Reservoir A).

Since the reservoir contains layers of shale, one of the possibilities for this behavior could be gas desorption below some critical pressure value. The available methane adsorption isotherms also duplicate the same behavior in the reservoir. In this example, the goal is finding the global properties over four units of that reservoir using the proposed methodology. The key parameters that control the global fluid flow behavior in the reservoir are reservoir compartments, aquifer strength, amount of shale gas contributing to the production, and the fracture intensity distributions. Reservoir B is in the same geological environment as Reservoir A but the current reservoir pressure is higher. The goal of history matching for this reservoir is finding the reservoir compartments, aquifer strength, and the fracture intensity distributions over one of the reservoir units.

5.4.2. Large-Scale Connectivity Estimation (Reservoir A)

The coarse-scale reservoir model in this example has four layers with each layer having 95 cells. The original geological model contained 49,633,080 grid blocks with 41,222,720 active cells. The facies map for this reservoir is depicted in **Fig. 5.4(a)**. 58

RFT pressure data were found for this reservoir and the fine model were up-scaled using the method described in Section 5.3 (Fig. 5.4(b)).



(a) (b)

Fig. 5.4 – (a) Facies map of all layers of Reservoir A, (b) up-scaled model for pressure matching step, colors represent grid index.

This coarse model provides a 3- to 4-order-of-magnitude decrease in reservoir simulation CPU time. Since flow-based upscaling is used to transfer the transmissibility from fine to coarse models, the resulting coarse model provides a good approximation to the pressure behavior in the fine model. More importantly, including many unknowns in the history matching problem leads to a severely ill-posed inverse problem. Hence, many possible solutions can be found that match the production history but provide distinct production forecasts. Since in the proposed method the control points are the same as the locations of available RFT data, higher grid resolution are located where the spatial

concentration of static pressure measurements is higher. Since the simulation CPU time for the coarse grid block system is about 15 to 20 seconds, a large number of samples are used for EnKF implementation to integrate the pressure data into the coarse-scale reservoir model. For the initial models, the transmissibility between the coarse grid blocks is calculated using the fine-scale geological model. Then, during the history matching process, the transmissibility multipliers and the aquifer strength are updated.

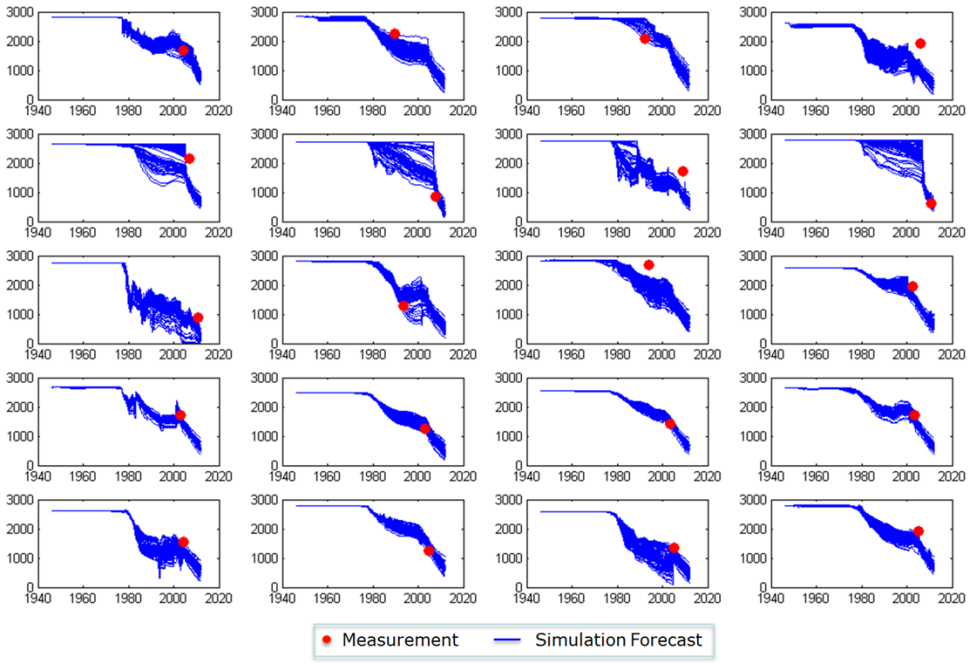


Fig. 5.5 – Final pressure forecast obtained from all the tuned coarse scale reservoir models.

Another problem for simulation of this reservoir is considering the gas desorption process. To account for gas desorption, one extra grid to each layer is added and connected to all the grids in that layer. The transmissibility between the added grids and their pore volumes provided additional history matching parameters. In the Petrel model, the reservoir has no fractures or flow barriers. Static pressure data is used to determine the higher/lower capability to flow inside the reservoir. The final result of pressure history matching is illustrated in **Fig. 5.5**.

To separate fractured, matrix dominated, and flow barriers, thresholds are applied on the final updated transmissibility multipliers. The transmissibility multipliers greater than 5 and less than 0.1 are assumed to represent fractured and flow barriers. The transmissibility multipliers between 0.1 and 5 are considered as matrix-dominated areas. The initial model does not have fracture and/or fault information; therefore, the initial model is matrix-dominated everywhere. Applying the threshold to the updated transmissibility multiplier ensemble generates a discrete distribution for transmissibility multipliers. **Fig. 5.6(a)** shows the mode of the discrete transmissibility multiplier distribution at each point.

In this figure, the fracture and matrix dominated regions as well as flow barriers are illustrated with green, blue, and red lines, respectively. Also, **Fig. 5.6(b)** illustrates the confidence map for the final transmissibility multipliers. The confidence map in Fig. 5.6(b) shows high confidence in the location of the fracture and two flow barriers in the northeast and southeast of the model. The final results don't show much confidence in the blue area in the west region of the transmissibility map (Fig. 5.6(a)). The low

transmissibility areas (red lines) in the final transmissibility multiplier map have good correlation with the fault tracks in the old Petrel model (Fig. 5.6(c)). Also, the high fracture intensity areas (green lines) show good correlation with high Gaussian curvature areas in Fig. 5.6(d).

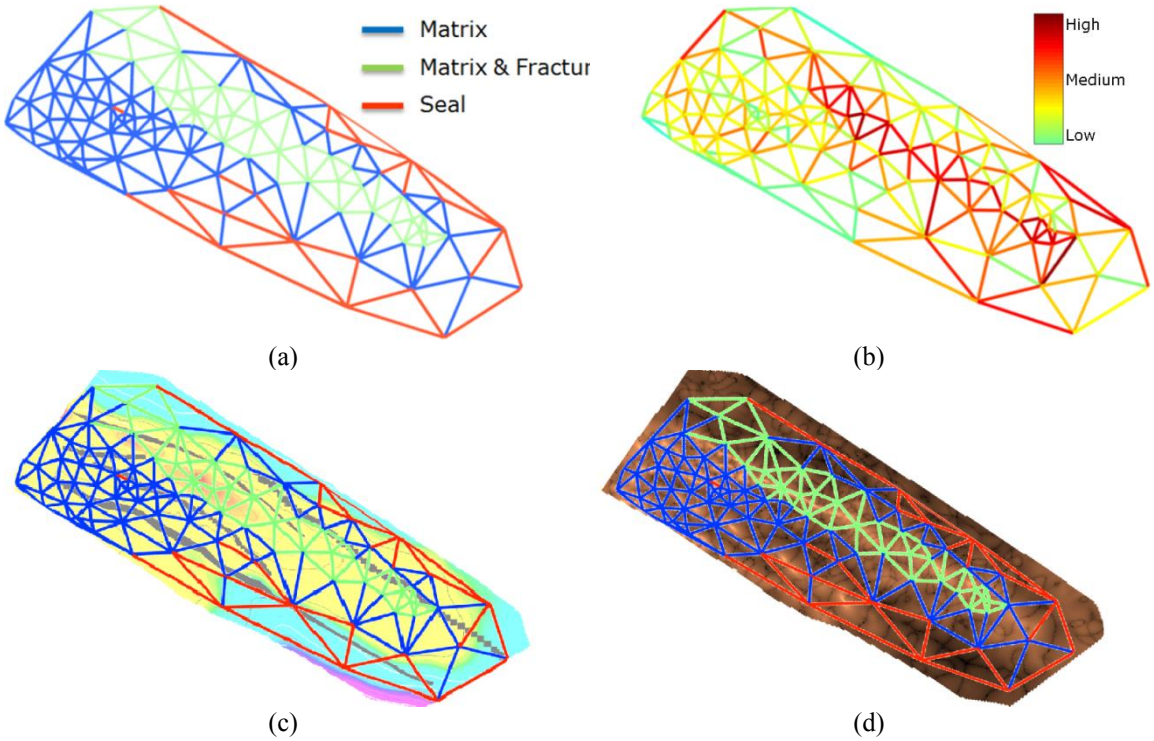


Fig 5.6 – (a) Final transmissibility multiplier map for one layer, (b) confidence map for final transmissibility multiplier map obtained from EnKF, (c) final transmissibility multiplier map for one layer on top of available fault traces in old Petrel model, (d) final transmissibility multiplier map for one layer on top of Gaussian curvature map.

Geologists have been using curvature analysis in the petroleum industry for four decades to identify the areas with greater density of fractures (Murray, 1965; Lisle, 1994; Ouenes et al. 1994; Ouenes et al. 1995; Blumentritt et al. 2006). Other tuning parameters to match the static pressure observations are aquifer pore volume and transmissibility multipliers associated with gas desorption process.

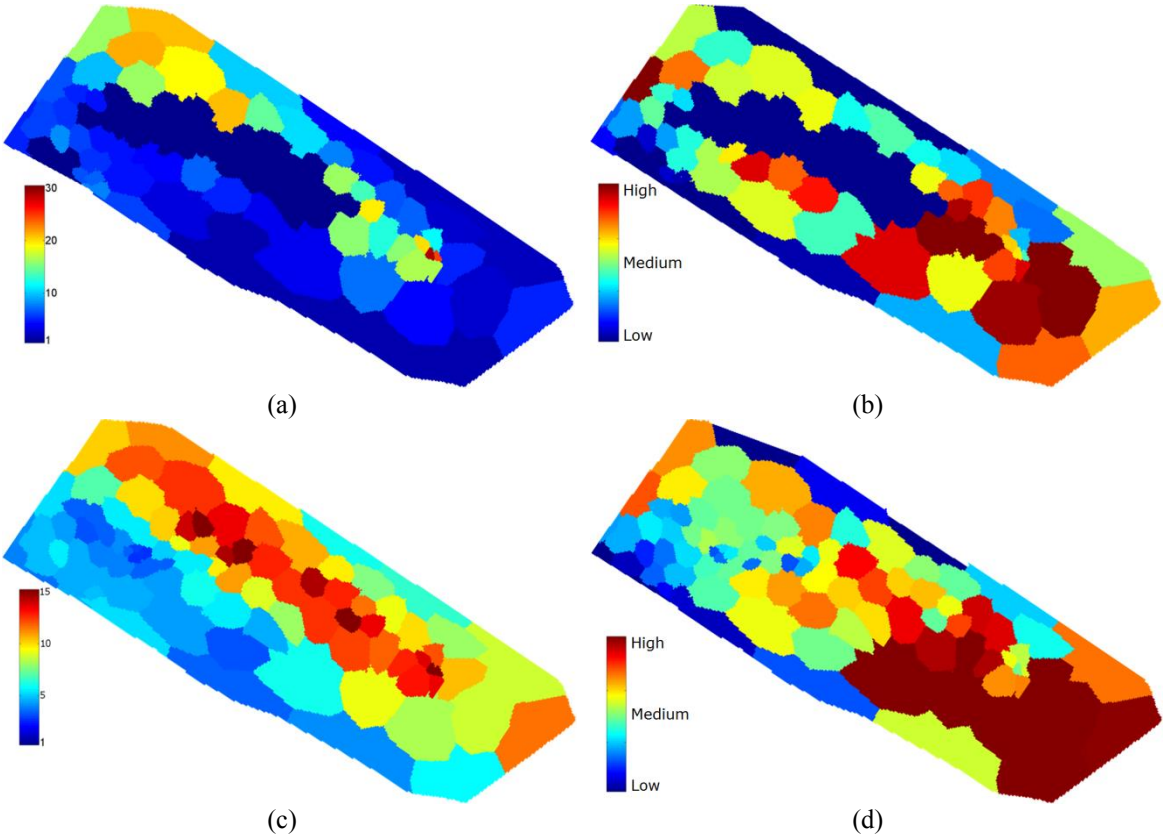


Fig 5.7 – (a) Final aquifer multiplier map for one layer, (b) confidence map associated with final aquifer multiplier map obtained from final ensemble (hotter values indicate higher confidence), (c) final transmissibility multiplier map and (d) confidence map associated with desorbed gas process.

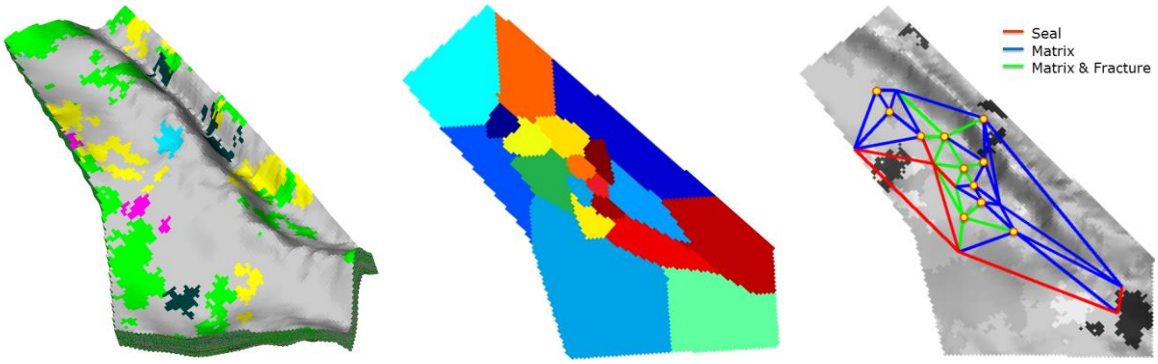
Fig. 5.7(a) shows the areal map of the aquifer pore volume multipliers. It illustrates that the aquifer has higher strength on the northern part of the reservoir than any other area. The dark blue area in the center of the map shows the regions initially above the water/oil contact; no aquifer grid blocks were considered for that section of the reservoir.

The confidence map, Fig. 5.7(b), shows relatively good confidence in the final aquifer multiplier values. Fig. 5.7(c) depicts the transmissibility multiplier map between each coarse grid block and the added grid block related to the simulation of gas desorption process. As illustrated in Fig. 5.7(c), high transmissibility multiplier areas have good correlation with high fracture intensity areas in Fig. 5.6(a). This result is consistent with the existing geological evidence. The third layer of the reservoir consists of shale and the desorbed gas could come from this layer. In high-fracture-intensity areas, reservoir layers have a higher chance of being connected to the shale layer.

5.4.3. Large-Scale Connectivity Estimation (Reservoir B)

The coarse-scale reservoir model in this example has one layer with 19 cells. The original geological model contained 3,902,892 active cells. The facies map for this reservoir is depicted in **Fig. 5.8(a)**. 11 RFT pressure data are found for this reservoir and upscaled the fine model using the method described in Section 5.3 (Fig. 5.8(b)). For the initial models, the transmissibility between the coarse grid blocks is calculated using the fine-scale geological model. Then, during the history matching process, the transmissibility multipliers and the aquifer strength are updated. As in the previous field study, no fractures or flow barriers are assumed in the initial reservoir model. The final

result of pressure history matching is illustrated in Fig. 5.8(c). The similar thresholds as previous example are used to separate fractured, matrix-dominated, and flow barrier areas. In Fig. 5.8(c), the fracture and matrix-dominated regions as well as flow barriers are illustrated with green, blue, and red lines, respectively.



(a) (b) (c)

Fig. 5.8 – (a) Facies map of reservoir B, (b) upscaled model for pressure matching step (colors represent grid index), (c) final transmissibility multiplier map and location of RFT measurements (yellow circles).

Based on the pressure history matching results, two compartments are detectable in the reservoir. In the primary dataset, all the RFT measurements are located in the north compartment. More investigation reveals that two more RFTs are available for the south part of the reservoir. The estimated pressure difference between the north and south parts of the reservoir, based on all RFT measurements, is close to 1000 psi, which is an indication of an actual flow barrier between these two compartments. Also, similar

to the previous example, the high fracture intensity areas (green lines) show good correlation with the high Gaussian curvature areas of the reservoir.

5.5. Conclusions

In this section, a novel approach is proposed for identification of field-scale reservoir connectivity by reconciling the low resolution of dynamic field pressure data and high resolution of geologic models. The proposed workflow is developed to convert high-resolution initial geologic models to low resolution (coarse-scale) grid systems for reservoir simulation. To reconcile data and model resolution, the grid system is generated using Delaunay triangulation by considering the location of the static pressure measurements as control points and distributing the unstructured grid blocks according to the spatial resolution of observations. The reservoir properties are upscaled from the fine-scale geological model to this coarse-scale unstructured grid system to create an initial static simulation model. The EnKF is applied to integrate the available data and automatically adjust the global model parameters including aquifer strength, global continuity/discontinuity of reservoir properties, and fault transmissibilities to match the static pressure.

After generating the coarse grid system, its performance to predict the static pressure measurement is evaluated using the fine-scale model. Then the proposed approach is combined with the EnKF to estimate important global parameters for two real naturally fractured reservoirs. The proposed approach provides promising history matching results upon application to the real field data. It provides a better

understanding of the overall continuity/discontinuity inside the reservoir model and identifies the high-fracture-intensity areas. Finally, comparison of the results with the existing geological evidence is used to support the validity of the estimation outcomes.

In the next section, the proposed workflow is applied in this section for another real field data set to infer the large scale fluvial channel continuity using the available pressure data. Then the finding in coarse scale history matching is downscaled to fine scale model and probability conditioning method is applied to conclude the locations of channelized fluvial.

6. WORKFLOW APPLICATION TO A FIELD CASE STUDY

6.1. Summary

A multi scale multi physics history matching approach is applied as the first step of a history matching workflow to infer the global reservoir continuity of an actual field. Since pressure variation represents a smooth function, an extremely low resolution (coarse scale) grid system is adopted for reservoir simulation. The grid system is generated through Delaunay triangulation by using the location of the static pressure measurements as control points and the unstructured grid blocks are distributed according to the spatial resolution of the observations. The global parameters such as aquifer strength and global continuity/discontinuity of reservoir properties are adjusted to match the static pressure. In the second step of the workflow, a fine-scale probabilistic interpretation of facies distribution is inferred using coarse-scale history matching results. In the final step, the probability condition method (PCM)-assisted ensemble Kalman filter (EnKF) approach is adopted to invert the flow data to update probabilistic spatial description of facies distribution (a probability map) and the resulting facies probability map is used to guide multipoint statistical (MPS) facies simulation from a specified training image (TI).

The field case data is obtained from an offshore turbidite reservoir. The fine-scale simulation model has almost 200,000 grid blocks. The reservoir has seven producers with two water injectors and over 5 years of production history including three-phase production history and static pressure data. The multi scale multi physics

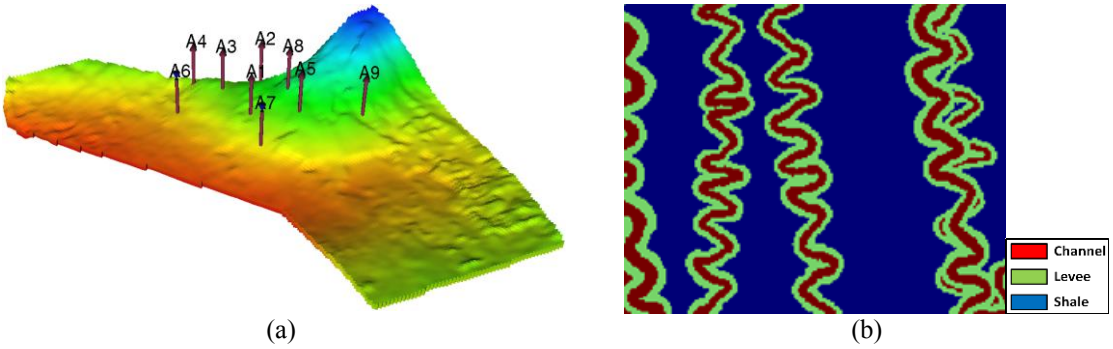
effectively detect the aquifer strength and overall continuity/discontinuity inside the reservoir using static pressure data. The coarse-scale history matching results are refined to provide the probabilistic interpretation about the location of channelized features. Using the PCM approach, geological continuities are preserved while history matched reservoir models are generated.

6.2. Reservoir Descriptions

The reservoir structure is characterized as a sand-filled depositional channel system and associated overbank in the presence of structural and stratigraphic traps. The reservoir forms a gentle anticline toward the southwest. A major fault elongates from east to west and provides a northern sealing boundary for the reservoir. The initial reservoir pressure is above saturation pressure, and the reservoir contains oil with dissolved gas. **Fig. 6.1(a)** illustrates the structural view of the reservoir along with the well locations. Seven production wells (A1to A5, A8, and A9) produce oil and associated gas from this reservoir while water is injected in two injectors (A6 and A7). Because of the asphaltene precipitation below the bubble point, the pressure maintenance was started almost a year after first day of production. Also, the annual static pressure data are available for almost all the producers along with the monthly production and injection rate measurement for over 5 years of operation.

The plot of log-permeability vs. porosity obtained from core data provides two distinguishable clusters. One cluster has low-permeability/low-porosity values, which represent shale facies, and the other one has high-permeability/high-porosity values,

which characterize by sand facies. Porosity ranges from 14% to 34%. Permeability shows a wide range, from 0.1 to 3000 md, and has a strong facies control. The average permeability for shale and sand facies are 0.4 and 700 md, respectively. Also, the reservoir contains levee facies with permeability of 200 md.



(a) (b)

Fig. 6.1 – (a) Structural view of the investigated reservoir with the production wells (A1-A5, A8, and A9) and injection wells (A6 and A7). (b) Training image contains meandering features used for sample generation.

Since the reservoir contains meandering channelized features, **Fig. 6.1(a)** is assumed to be the possible training image for the reservoir. 50 samples are generated from that training image. Fig. 6.2 illustrates the forecast of static pressure and field water cut using initial samples. The red dots represent the actual measurements. The simulated pressure falls rapidly and the simulated field water cut is lower than the field measurements.

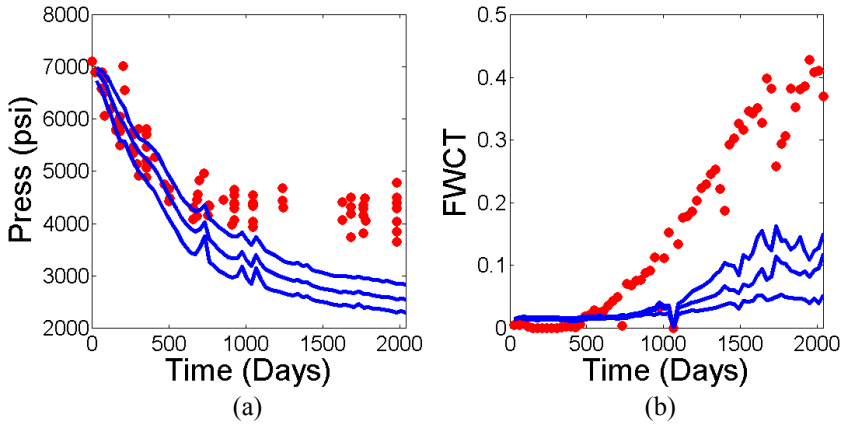


Fig. 6.2 – Actual observation (red dots) and quantile plot of (a) field average reservoir pressure and (b) field water cut using initial 50 samples.

6.3. Global Reservoir Connectivity Parameters

The fine-scale simulation model was generated using available seismic and log data. The fine-scale model has 194,250 grid blocks with 95 static pressure data gathered during the first 5 years of production. The upscaled reservoir model using the approach described in Section 5.3 is depicted in **Fig. 6.3**. The coarse scale model has 38 grid blocks. Since flow-based upscaling was used to transfer the transmissibility from the fine to the coarse model, the resulting coarse model provides a good approximation to the pressure behavior in the fine model. More importantly, including many unknowns in the history matching problem leads to a severely ill-posed inverse problem. Hence, many possible solutions can be found that match the production history but provide distinct production forecasts. Since in the proposed method the control points are the same as the locations of available static pressure data, higher grid resolution is located where the spatial concentration of static pressure measurements is higher.

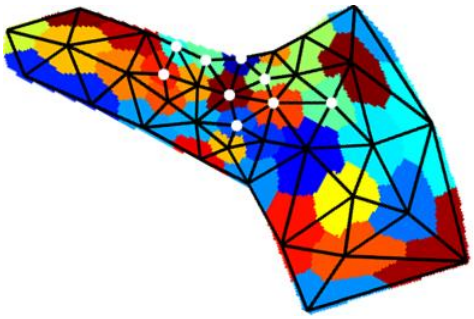


Fig. 6.3 – Coarse-scale reservoir model according to the well locations.

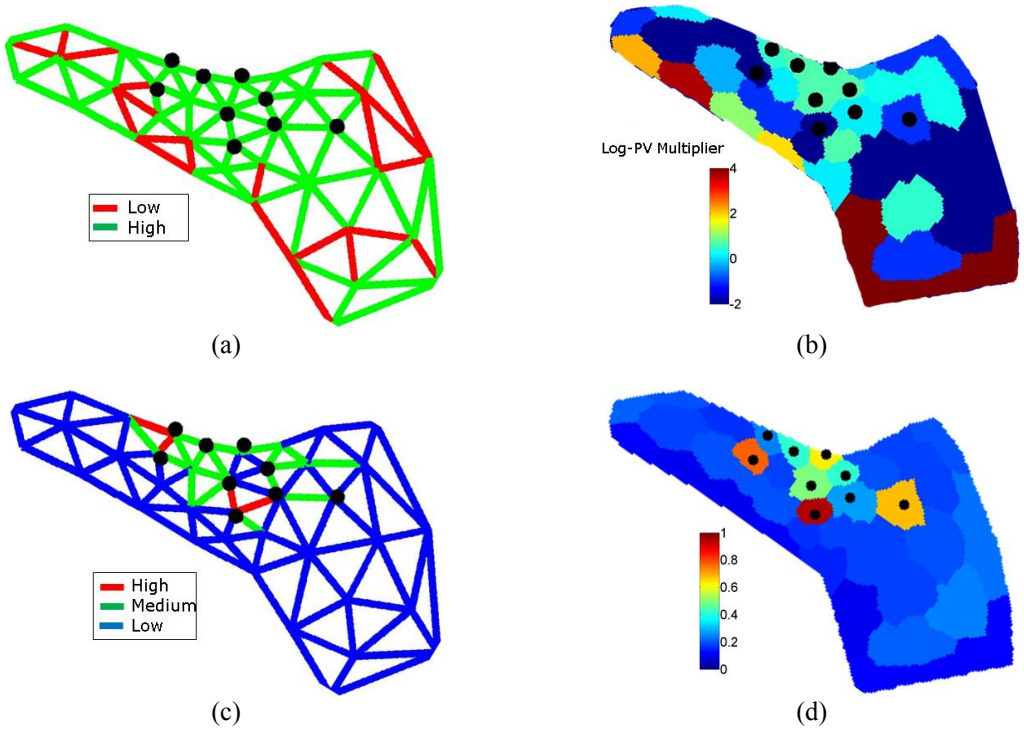


Fig. 6.4 – Final history matched maps of (a) transmissibility multipliers and (b) pore volume multipliers. The confidence maps of (c) transmissibility multipliers and (d) pore volume multipliers.

In the first step of data assimilation, the transmissibility between the coarse grid blocks is calculated using the fine-scale geological model. Then, during the coarse-scale static pressure history matching process, the transmissibility multipliers and the pore volume multipliers are updated. The history matched pore volume of those grid blocks located below the water/oil contact represent the energy support by the aquifer. The results of static pressure history matching are illustrated in **Fig. 6.4**.

As illustrated in Fig. 6.4(a), there is good connectivity between the wells and some discontinuities on the east, west, and south of the reservoir. Fig. 6.4(b) illustrates the log-pore volume multiplier map. The pore volume map demonstrates that the aquifer has more strength on the south side of the reservoir. Also, Fig. 6.4(c) and (d) show higher confidence near the wells and less confidence away from the wells.

6.4. Downscale Pressure History Matched Model

After history matching the coarse-scale model, the learning from the first step of proposed workflow must be downscaled to fine-scale and then tuned for the water-cut history matching. The concept of soft data integration, Subsection 1.3.1, is used to transfer the information from the coarse-scale to the fine-scale. Before presenting the results of downscaling for the investigated field, the procedure using a synthetic model is presented.

The 2D model in **Fig. 6.5(a)** is used as the true synthetic model. The model contains channelized features which connect in the middle of the model. The training image that is consistent with the synthetic model is depicted in Fig. 6.5(b). I assigned

zero and one to the channel and background facies, respectively. $3 \times 3, 5 \times 5$, and 15×15 upscaling schemes are used to upscale assigned values. Arithmetic averaging is used as the upscaling approach. The upscaled 2D maps are presented in Fig 6.5(c). Then, 100 samples are generated using a Snesim algorithm and condition on the upscaled maps. Also, for comparison, 100 samples are generated without conditioning on any soft data. The mean of the 100 samples using different soft data along with one sample from each set is presented in Fig. 6.5(d) and (e). In Fig. 6.5, the second row illustrates sampling without conditioning on any soft data; the mean is clearly different from other cases. As the resolution of the upscaled map decreased, the similarity of the generated samples to the original model decreased.

Even though the least amount of information is carried on the probability map using the 15×15 upscaling scheme, still the generated samples show the main features of the original model. More importantly, the sharp edges of probability maps are removed by MPS sampling. In the actual field problem, the upscaled field description is obtained from static pressure history matching.

Fig. 6.6(a) shows the probability map obtained from the results of static pressure history matching. 50 samples are drawn from the designed training image (Fig. 6.1(a)) and conditioned on the estimated probability map from the first step of history matching using the Snesim algorithm. The generated fine-scale models were used as the initial model for the water-cut history matching step. The mean of the initial fine-scale model is depicted in Fig. 6.6(b).

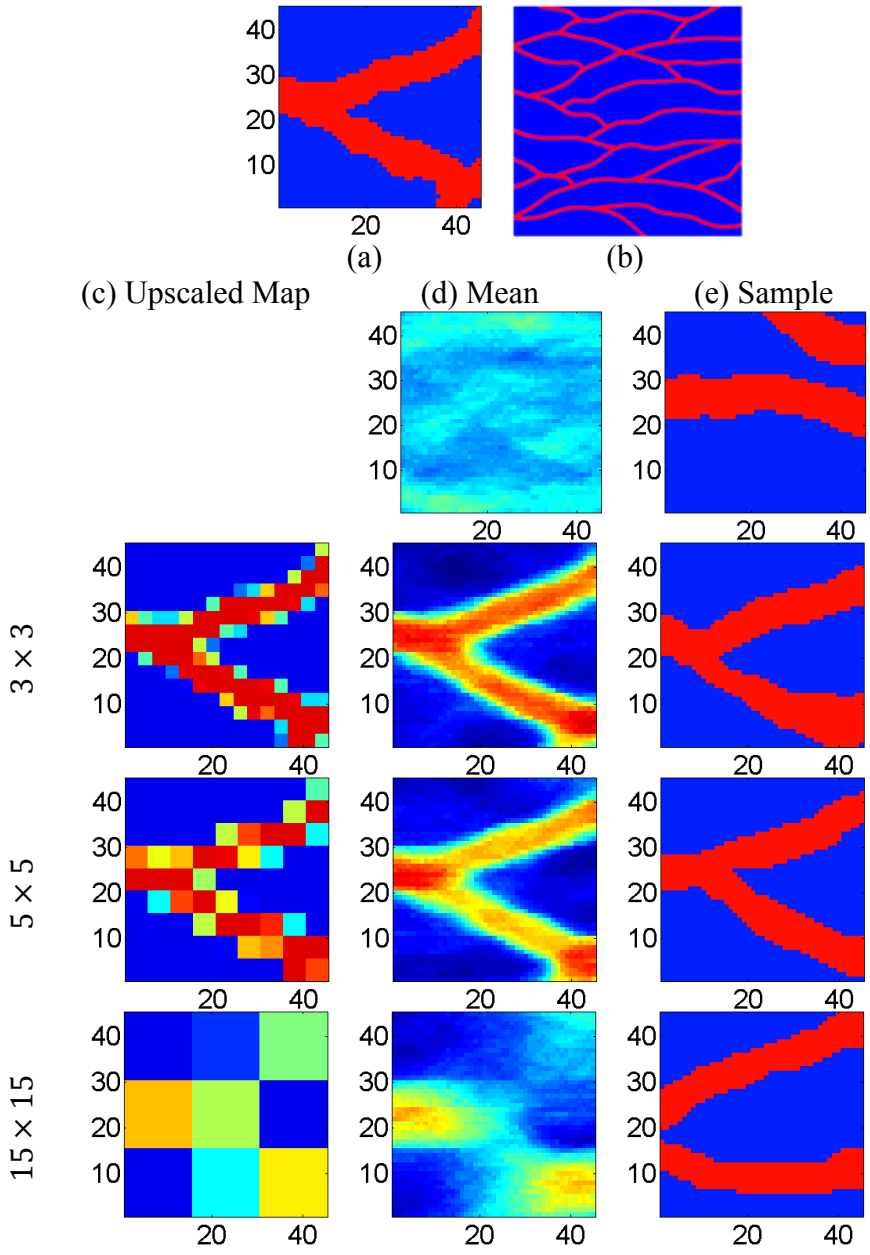
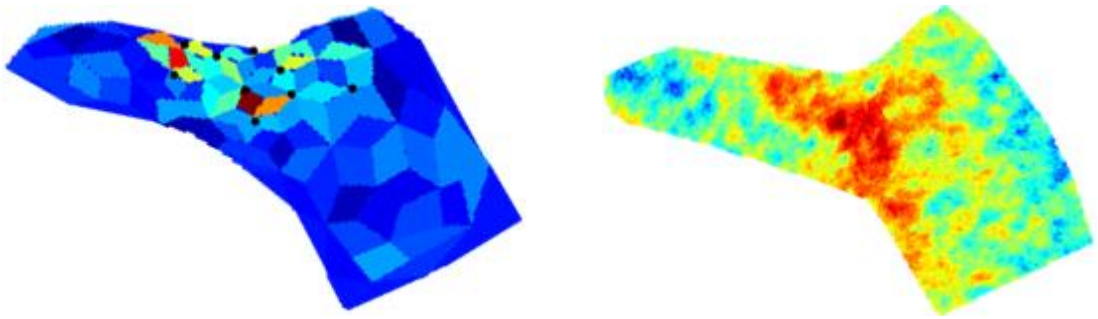
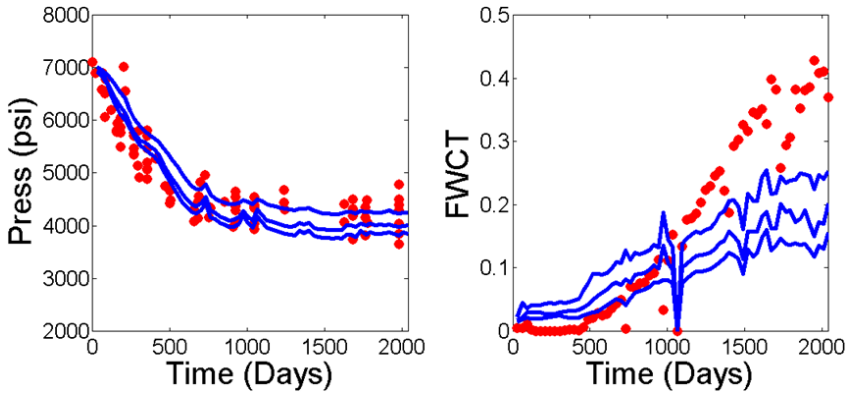


Fig. 6.5 – (a) True synthetic model and (b) training image used for transfer learning example from coarse model to fine model. (c) Upscaled map obtained from different upscaling schemes, (d) mean of 100 samples, and (e) one sample using the Snesim approach and conditioned on the probability map.



(a) (b)

Fig. 6.6 – (a) Probability map generated from static pressure history matching and (b) mean of 50 samples generated from the training image in Fig. 6.1(a) and conditioned on the probability map in (a).



(a) (b)

Fig. 6.7 – Actual observation (red dots) and quantile plot of (a) field average reservoir pressure and (b) field water cut using initial fine-scale reservoir models.

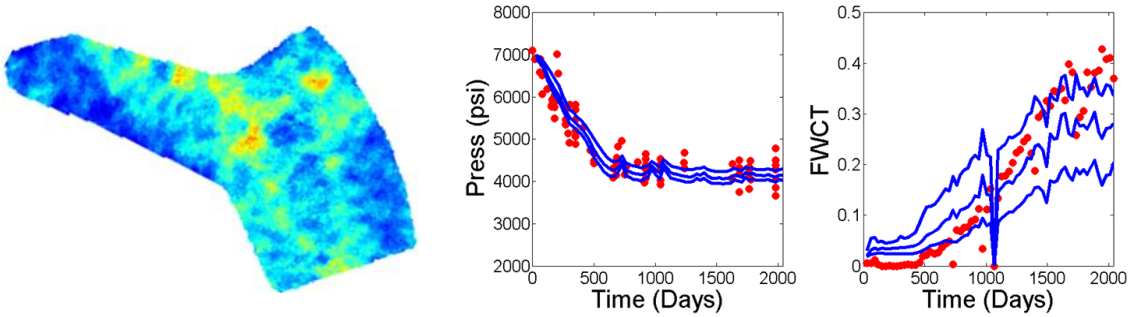
Forecast is obtained using the initial fine models. The field average pressure and field water-cut forecast are illustrated in **Fig. 6.7**. The field average pressure prediction (Fig. 6.7(a)) is improved in comparison with the forecast obtained from initial samples

generated from the training image (Fig. 6.2(a)). Even though matching the water cut is not the target for the first step of history matching, the field water cut (Fig. 6.7(b)) shows improvement, too. The simulation models provide less water production in the later time; therefore, the goal in the second step of history matching is matching the field water cut.

6.5. Fine-Scale Water-Cut History Matching

To match the field water cut, the PCM, which is presented in Section 2, is applied. The initial ensemble of facies generated in the previous subsection is used as the initial models for fine-scale history matching. Also, the probability maps obtained from coarse-scale history matching were used as the initial probability maps. After assigning permeability values to each facies, the facies ensemble is used to predict the observed fluid flow quantities for the next time step. In the analysis step, the log-permeability realizations that were used to find the log-permeability mean and the probability map are updated. The update equation requires up to the second-order statistics of the estimation parameters. The ensemble of log-permeability fields is used to derive the first- and second-order statistics required during the update step.

The updated mean of the ensemble was then used to infer probabilistic information about facies distribution. I assumed that the probability of each facies at a given grid block is a function of the distance between the updated log-permeability value at that grid block and the log-permeability value of each facies. Only the analysis equation is used to estimate the log-permeability mean from the production data.



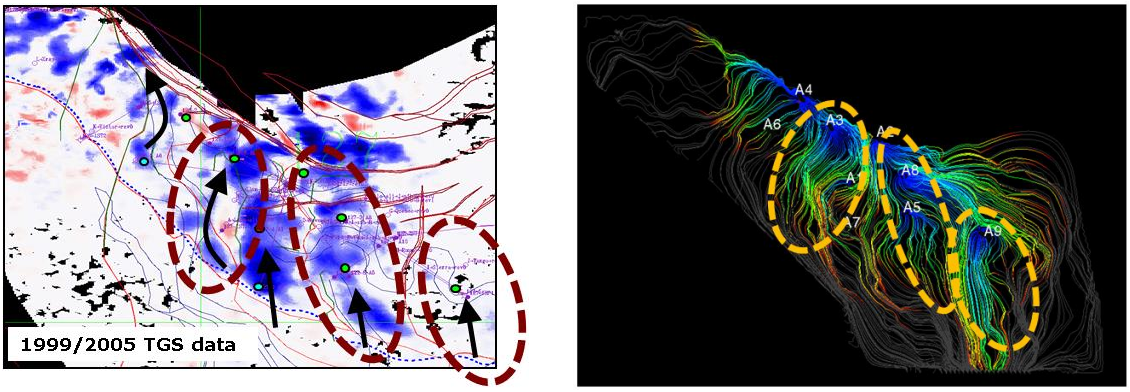
(a) (b) (c)

Fig. 6.8 – (a) Final probability map obtained from fine-scale history matching and actual observation (red dots) and quantile plot of (b) field average reservoir pressure and (c) field water cut using final fine-scale reservoir models.

The final updated channel probability map along with the final field average pressure and field water cut are depicted in **Fig 6.8**. Even though the initial and updated probability maps (Fig. 6.8(a)) show similar trends around the well locations, the main improvement in the final probability happened in the south part of the reservoir. The final probability map shows channel connectivity from the aquifer to the well location which is not significant in the initial probability map. The final field water cut (Fig. 6.8(b)) forecast confirms the same behavior. Also, the final forecasts of average field pressure (Fig. 6.8(c)) are slightly improved in comparison with the previous step.

4D seismic is another source of information available for the investigated field. The actual water flow paths from the aquifer to the reservoir are depicted in **Fig. 6.9(a)**. To make a comparison, the best history matched reservoir model is combined with streamline simulation to predict the flow paths inside the reservoir. The time of flights

depicted in Fig. 6.9(b) obtained from the streamline simulation show a similar trend to the actual flow paths inside the reservoir.



(a) (b)

Fig. 6.9 – (a) Actual and (b) predicted flow paths for the investigated reservoir.

6.6. Conclusions

In this section, the proposed workflow is used to integrate static pressure and water cut measurement into the reservoir models for an actual field in deepwater Gulf of Mexico. The novel approach described in Section 5 is used to identify the field-scale reservoir connectivity from static pressure measurements. The coarse scale grid system is generated from the well locations. By tuning the coarse grid transmissibility multipliers and aquifer size, the actual average field pressure measurements are predicted. Also, the field water-cut prediction is improved. In the second step, the learning from the coarse-

scale history matching is transferred to the fine-scale reservoir model through the probability map. Finally, the probability conditioning method (PCM), which is presented in Section 2, is applied to match the water cut by updating the probability of existing features inside the reservoir. The final history matched reservoir model can correctly predict the field's average pressure behavior and water cut. Also, the time of flights obtained from the streamline simulation and available 4D seismic map show similar water flow paths.

7. DISSERTATION CONTRIBUTION AND FINAL CONCLUSIONS

In this dissertation, three new techniques are described to preserve and identify large-scale geological features in reservoirs during history matching. In Section 2, a probability conditioning method (PCM) was developed for constraining multipoint statistical (MPS) facies simulation to nonlinear flow data. Implementing the PCM with the ensemble Kalman filter (EnKF) results in an improved performance of the filter updates through the preservation of the facies correlation structure and the introduction of additional ensemble variability (spread) due to the resampling of facies from the training image (TI) after each update step.

As an extension to Section 2, in Section 3 an adaptive strategy is presented through a data feedback mechanism based on the predictive performance of the past realizations (sampling history) to improve the conditioning efficiency. The method is presented as a stochastic optimization algorithm and extended to the case where multiple training images are used as alternative plausible interpretations of the geologic scenarios for a given formation.

In Section 4, a novel Bayesian mixture model for adaptive and efficient sampling of conditional facies from multiple uncertain TIs was developed. The posterior distribution of facies is partitioned into individual conditional densities of the TIs and the corresponding mixture weight from the likelihood function for each TI is estimated. To implement the conditional sampling, a recently developed EnKF-based probability conditioning method is adopted, whereby EnKF is used to invert the flow data and

obtain a facies probability map (soft data) to guide conditional facies simulation from each TI.

In Section 5, a new workflow for integration of pressure data for estimating large-scale reservoir connectivity was described. According to the spatial resolution of the observations, the unstructured grid system is generated through Delaunay triangulation by using the location of the static pressure measurements as control points. The important advantages of this workflow for characterization of field-scale reservoir connectivity from pressure data include very fast connectivity estimation with a low-order model and effective parameterization to reduce the number of unknowns to a level commensurate with the available static pressure measurements.

At the end of Section 5 and in Section 6, the workflow from Section 5 is applied to three real field data sets to infer the large-scale feature continuity using the available pressure data. Then, in Section 6, the findings from coarse-scale history matching is downscaled to a fine-scale model and applied a probability conditioning method to conclude the locations of fluvial channels.

The main results and findings from Sections 2 to 6 are summarized in Subsection 7.1. Subsection 7.2 concludes with the description of future work within the context of preserving the large-scale feature continuity during reservoir history matching.

7.1. Dissertation Contributions and Conclusions

The dissertation starts with presenting a new probability conditioning method (PCM) for constraining TI-based facies simulation to reproduce dynamic flow measurements.

While the EnKF update equation was employed for production data integration, the PCM approach is quite general and can be implemented by means of alternative inverse modeling techniques. Since discrete geologic facies contain features that are more amenable to description with categorical random variables, standard grid block property estimation with the EnKF can lead to inconsistent facies connectivity and spurious updates, especially at distant locations from observations. Although EnKF has been shown to be an effective history matching method, the continuous, second-order, and unconstrained nature of its update equation is poorly suited to the estimation of discrete geologic objects (such as fluvial channels) which can only be accurately described if higher-order statistical moments are included. The proposed probability conditioning method uses the EnKF to update the map of the log-permeability mean from the production data (and a prior ensemble of models). This is then used to infer a probabilistic description of facies distribution. The resulting probability map is incorporated into TI-based multipoint geostatistical simulation algorithms to generate conditional facies realizations.

Facies resampling (from the TI) after each EnKF update step offers several advantages over the standard implementation of EnKF. Some of these advantages include the following: (1) consistency of the updated facies with the prior TI (geologic model) and preservation of higher-order statistics; (2) probabilistic interpretation of the production data to reduce the effect of spurious EnKF updates; and (3) improvement of the ensemble variance underestimation (and collapse) issues through the introduction of

additional ensemble variability, which is done by resampling new facies from the TI after each update step.

In Section 2, the implementation of the proposed framework was presented and discussed, along with synthetic two-dimensional waterflooding examples with different levels of TI representation. This showed that the proposed framework could effectively incorporate the production data into conditional facies simulation using the Snesim algorithm. The production forecasts that were obtained when the proposed approach was used provided wider ensemble spread than those obtained from the estimation of grid-based log-permeability maps, mainly because of the variability and randomness introduced by resampling new facies from the TI. Additional complex examples reveal that the method could not reproduce the correct facies orientation and structure when the TI was completely inconsistent with the true model.

An adaptive sampling approach for conditioning MPS-based facies simulation from a TI to nonlinear flow data is presented in Section 3. This is achieved by storing and using previously accepted realizations to construct a probability map that describes the likelihood of facies occurrence in each grid block. To increase the sampling efficiency, an initially uninformative probability map is gradually updated by incorporating the information about the facies distribution in the chain of accepted realizations. Since the spatial facies patterns in the accepted realizations reproduce the observed data (in a probabilistic sense), the dominant facies connectivity in them (common features among the accepted realizations) are likely to represent some of the existing trends in the solution. Hence, a collective probabilistic representation of the

common features in the accepted realizations is used to guide and accelerate the sampling process.

The implementation of the proposed approach is illustrated its performance using two set of numerical waterflooding experiments, one using a single given TI and the other using multiple uncertain TIs. The proposed algorithm was able to effectively sample from the given TIs as conceptual prior models while incorporating the observed data through the statistical information contained in the accepted realizations. Feeding the observed data back into the simulation process through a facies probability map provided an effective mechanism to guide the simulation of the new facies realizations and improved the sampling efficiency.

Conventional metrics of performance, such as computation and accuracy, obtained from the non-adaptive algorithm and adaptive implementation demonstrate that the adaptive method offers considerable improvements over the non-adaptive sampling. The proposed adaptive approach generates more representative realizations in fewer iterations and leads to significant computational savings in practical applications. A version of the proposed algorithm for application under multiple Tis is presented. The extension of the adaptive algorithm for use under TI uncertainty could distinguish between consistent and inconsistent TIs.

Bayesian mixture model is presented for adaptively conditioning the simulation of geologic facies from multiple TIs to nonlinear pressure and flow measurements in Section 4. The mixture model approach uses the PCM data integration algorithm to adaptively draw conditional facies realizations from multiple TIs. The PCM was applied

to convert dynamic pressure and flow data into facies probability map, which, in turn, was used to guide facies simulation from prior TIs. At each sampling stage, the TI weights were estimated based on the likelihood function for the individual prior TIs and determined the number of samples drawn from each.

After presenting the implementation details of the proposed approach, its performance is examined under several prior assumptions, including uncertainty in formation connectivity (2D and 3D models) and channel direction (orientation) for a given TI in fluvial systems. In all cases, the method was eventually able to identify, and accordingly sample from, the correct TI. Also, the efficiency of the proposed approach is investigated for a two-phase 3D problem under structural uncertainty. The optimized well controls estimated using the best MMPCM model and true synthetic model provide similar net present values (NPVs), while optimized control estimated using the best standard EnKF results in 10% less NPV. More importantly, the similar unswept area in the oil saturation maps obtained from the best MMPCM and true synthetic model offer a good candidate for infill drilling. This unswept area is missing in the saturation prediction using the best EnKF model.

Although the pattern-imitating nature of MPS simulation from prior TIs presents an opportunity to model more complex geologic phenomena, it also poses an important risk when the prior TI fails to represent the correct facies connectivity. Since the resulting facies models often dominate fluid displacement behavior, the TI uncertainty must be taken into consideration in MPS simulation. Conditioning the MPS simulation

results on nonlinear dynamic data remains an important topic in application of this method to modeling complex subsurface systems.

The complexity of the conditional simulation problem increases when the uncertainty in the TI model has to be acknowledged and incorporated. When prior knowledge about the structural connectivity is not adequate to overwhelmingly support the use of a single TI, a wider range of possible structural connectivity models (TIs) should be considered. Evaluations should also rely on the dynamic flow data to distinguish between alternative TI candidates and adaptively sample from them based on their predictive performance (likelihood to reproduce the observed data).

In Section 5, a novel approach for identification of field-scale reservoir connectivity by reconciling the low resolution of dynamic field pressure data with the high resolution of geologic models is described. In the proposed workflow, an approach is developed to convert high-resolution initial geologic models to low-resolution (coarse-scale) grid systems for reservoir simulation.

To reconcile data and model resolution, the grid system is generated using Delaunay triangulation by considering the location of the static pressure measurements as control points and the unstructured grid blocks are distributed according to the spatial resolution of observations. The reservoir properties are upscaled from the fine-scale geological model to this coarse-scale unstructured grid system to create an initial static simulation model. The EnKF is applied to integrate the available data and it automatically adjusts the global model parameters including aquifer strength, global

continuity/discontinuity of reservoir properties, and fault transmissibilities to match the static pressure.

After generating the coarse grid system, its performance to predict the static pressure measurement is evaluated using the fine-scale model. The proposed approach is combined with the EnKF to estimate important global parameters for two real naturally fractured reservoirs. The proposed approach provides promising history matching results upon application to the real field data. It provides a better understanding of the overall continuity/discontinuity inside the reservoir model and identifies the high-fracture-intensity areas. Finally, comparison of the results with the existing geological evidence was used to support the validity of the estimation outcomes.

In the last section, a workflow is proposed to integrate static pressure and water-cut measurement into the reservoir models for an actual field in deepwater Gulf of Mexico. The novel approach described in Section 5 is applied to identify the field-scale reservoir connectivity from static pressure measurements. By tuning the coarse-grid transmissibility multipliers and aquifer size, the actual average field pressure measurements are predicted. Also, the field water-cut prediction is improved.

In the second step, the learning from the coarse-scale history matching is transferred to the fine-scale reservoir model through the probability map. Finally, the PCM is applied to match the water cut by updating the probability of existing features inside the reservoir. The final history matched reservoir model can correctly predict the field average pressure behavior and field water cut. Also, the time of flight obtained

from the streamline simulation and available 4D seismic map show similar water flow paths.

7.2. Future Research Directions

This subsection discusses the future potential for research direction. Even though the PCM and MMPCM approaches are successfully applied using EnKF and Snesim algorithms, many other data integration and MPS sampling methods can be used with the suggested algorithms. Each data integration and MPS sampling approach has its own pros and cons, and the PCM and MMPCM approaches might need to be modified. For example, a useful application could be combination of PCM and Markov chain Monte Carlo (MCMC) and Snesim algorithm. To perform this task, the major modification could happen in the software using the Snesim method. Since the current software that generates samples using the Snesim algorithm does not provide the probability of each generated sample, it is impossible to accurately calculate the transition distribution between two generated samples.

Another modification in the PCM approach could be linking the variance map obtained from EnKF and a tau value. In current implementation, the tau value is uniform for every grid block. The main reason for choosing a uniform tau value is the software limitation. In the pressure matching step using coarse-scale unstructured grid blocks, the number of unknowns was reduced to decrease the ill-posedness of the inverse problem. However, the large-scale geological continuity could elongate beyond the interwell distances. In this situation, unstructured parameterizations like the grid connectivity

transform (GCT) are useful to even further reduction of unknown parameters and the ill-posedness of the inverse problem.

REFERENCES

Albertoni, A., Lake, L.W. (2003), Inferring interwell connectivity only from well-rate fluctuations in water floods SPEREE 6(1), 6–16 (SPE-83381-PA).

Alcolea, A., and P. Renard (2010), Blocking moving window algorithm: conditioning multiple-point simulations to hydrogeological data, Water Resour. Res., 46, W08511, doi:10.1029/2009WR007943.

Arpat G, Caers J (2007), Conditional simulation with patterns. Math Geol 39(2):177–203

Aziz K., A. Settari (1979), Petroleum reservoir simulation, Springer.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J. (2001), Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. Mon. Weather Rev. 129:420–436.

Blazkova, S., and K. Beven (2009), A limits of acceptability approach to model evaluation and uncertainty estimation in flood frequency estimation by continuous simulation: Skalka catchment, Czech Republic, Water Resour. Res., 45, W00B16, doi:10.1029/2007WR006726.

Burgers, G., Leeuwen, P. J. van, and Evensen, G. (1998), Analysis scheme in the ensemble Kalman filter. Mon. Weather Rev., 126: 1719–1724.

Caers, J., and B. T. Hoffman (2006), The probability perturbation method: A new look at Bayesian inverse modeling, Math. Geol., 38(1), 81–100, doi:10.1007/s11004-005-9005-9.

Caers, J., and T. Zhang (2004), Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models, Grammer GM (ed) AAPG Memoir, Integration of Outcrop and Modern Analogs in Reservoir Modeling, 80:383–394.

Caers, J., S. Krishnan, Y. D. Wang, and T. A. Kovscek (2002), Geostatistical approach to streamline-based history matching, SPE Journal, 7, 250–266, doi:10.2118/73144-PA.

Carle, S. F., E. M. Labolle, G. S. Weissmann, D. Van Brocklin, and G. E. Fogg (1998), Conditional simulation of hydrofacies architecture: A transition probability/Markov approach, in hydrogeologic models of sedimentary aquifers, Concepts in Hydrogeology and Environmental Geology, vol. 1, 147–170, SEPM Publications, Tulsa, Okla.

Carrera, J. (1987), State of the art of the inverse problem applied to the flow and solute transport problems, Groundwater Flow and Quality Modeling, NATO ASI Ser., 549–585.

Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and L. Slooten (2005), Inverse problem in hydrogeology, Hydrogeol. J., 13(1), 206–222, doi:10.1007/s10040-004-0404-7.

Carrera, J., Neuman, S.P. (1986a), Estimation of aquifer parameters under transient and steady-state conditions: 1-Maximum likelihood method incorporating prior information, Water Resour. Res. 22(2), 199-210.

Carrera, J., Neuman, S.P. (1986b), Estimation of aquifer parameters under transient and steady-state conditions: 2-Uniqueness, stability and solution algorithms, Water Resour. Res. 22(2), 211-227.

Carrera, J., Neuman, S.P. (1986c), Estimation of aquifer parameters under transient and steady-state conditions, 3, Applications, Water Resour. Res. 22(2), 228-242.

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

Chen, Y., Durlofsky, L. J., Gerritsen, M. G. and Wen, X. H. (2003), A coupled local-global upscaling approach for simulating flow in highly heterogeneous formations, Advances in Water Rescources 26, 1041–1060.

Chiles JP, and P. Delfiner (1999), Geostatistics: modeling spatial uncertainty, Wiley, New York.

Cooley, R.L. (1979), A method of estimating parameters and assessing reliability for models of steady state groundwater flow: 2. Application of statistical analysis. Water Resources Research 15(3), 603–617.

de Marsily, G., Delay, F., Goncalves, J., Renard, P.H., Teles, V., and Violette, S. (2005), Dealing with spatial heterogeneity, Hydrogeology Journal 13(1), 161–183.

de Marsily, G., Delhomme, J., Delay, F., and Buoro, A. (1999), 40 years of inverse problems in hydrogeology, CR Acad Sci Ser II A Earth Planet Sci., 329(2),73–87.

de Marsily, G., Lavedan, G., Boucher, M., and Fasanino, G., (1984), Interpretation of interference tests in a well field using geostatistical techniques to fit the permeability distribution in a reservoir model, G. Verly, Editor, et al.Geostatistics for natural resources characterization. Part 2, D. Reidel Pub Co, pp. 831–849.

Deutsch, C., and A. G. Journel (1998), GSLIB: Geostatistical software library and user's guide, 2nd edition, 368 pp., Oxford University Press, Oxford, U.K.

Deutsch, C.V., and L.Wang (1996), Hierarchical object-based geostatistical modeling of fluvial reservoirs, paper SPE 36514 presented at the 1996 SPE Annual Technical Conference and Exhibition, Denver, Oct 6–9.

Doherty, J. (2003), Ground Water model calibration using pilot points and regularization, Groundwater, 41(2), 170-177.

Doherty, J., and B. E. Skahill (2006), An advanced regularization methodology for use in watershed model calibration, J. Hydrol., 327, 564–577, doi:10.1016/j.jhydrol.2005.11.058.

dos Santos Amorim E.P., P. Goldfeld, F. Dickstein, R.W. dos Santos, and C.R. Xavier (2010), Automatic history matching in petroleum reservoirs using the TSVD method, Comp. Sci. and Its Applications, 6017, 475-487

Durlofsky, L. J. (2005), Upscaling and gridding of fine scale geological models for flow simulation, Proceedings of the 8th International Forum on Reservoir Simulation, Iles Borromees, Stresa, Italy.

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

Evensen G (2004), Sampling strategies and square root analysis schemes for the EnKF. Ocean Dyn 54:539–560

Evensen, G. (1992), Using the extended Kalman filter with a multilayer quasi geostrophic ocean model. J. Geophys. Res., 97(C11):17905–17924.

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

Evensen, G. (2007), Data assimilation: the Ensemble Kalman Filter. New York: Springer

Farmer, C. L. (2002), Upscaling: A review, International Journal for Numerical Methods in Fluids 40, 63–78.

Fu, J., Gomez-Hernandez, J. J. (2009), A blocking Markov chain Monte Carlo method for inverse stochastic hydrogeological modeling. Mathematical Geosciences 41 (2), 105-128.

Gavalas, G.R., P.C. Shah, and J.H. Seinfeld (1976), Reservoir history matching by Bayesian estimation, SPEJ 16 (6), 337-350. SPE-5740-PA.

Gerritsen, M. G. and Durlofsky, L. J. (2005), Modeling fluid flow in oil reservoirs, Annual Review of Fluid Mechanics 37, 211–238.

Gomez-Hernandez, J. J., Wen, X. H., 1998. To be or not to be multi-Gaussian? A reflection on stochastic hydrogeology. Advances in Water Resources 21 (1), 47-61.

Gomez-Hernandez, J., Sahuquillo, A., and Capilla, J. (1997), Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data. 1. theory, Journal of Hydrology, 203:162–174.

Goovaerts, P. (1997), Geostatistics for natural resources evaluation, Oxford University Press, New York

Gotzinger, J., and A. Bardossy (2008), Generic error model for calibration and uncertainty estimation of hydrological models, Water Resour. Res., 44, W00B07, doi:10.1029/2007WR006691.

Guardiano, F., and R. Srivastava (1993), Multivariate geostatistics: Beyond bivariate moments, Geostatistics-Troia, Kluwer Academic Publications, Dordrecht, 1,133–144.

Haldorsen, H.H., and L.W. Lake (1984), A new approach to shale management in field-scale models, Soc. Pet. Eng. J., 24(8), 447–452

Heffer, K.J., Fox, R.J., McGill, C.A., Koutsabeloulis, N.C., (1997), Novel techniques show links between reservoir flow directionality, earth stress, fault structure, and geomechanical changes in mature waterfloods. SPEJ2(2),91–98.

Hendricks Franssen, H.J., Alcolea, A., Riva, M., Bakr, M., van de Wiel, N., Stauffer, F., and Guadagnini, A., (2009), A comparison of seven methods for the inverse modelling of groundwater flow. Application to the characterisation of well catchments. Advances in Water Resources, doi: 10.1016// j.advwatres. 2009.02.011.

Hill, M.C., and C.R., Tiedeman (2007), Effective groundwater model calibration, John Wiley & Sons, New York, NY.

Hoffman, B.T, and Caers, J. (2005), Regional probability perturbations for history matching, Journal of Petroleum Science and Engineering, 46, 53-71

Holden, L., R. Hauge, O. Skare, A. Skorstad (1998), Modeling of fluvial reservoirs with object models, Math. Geol., 30(5), 473–496

Honarkhah, M., and J. Caers (2010), Stochastic simulation of patterns using distance-based pattern modeling, Math Geos, 42(5), 487-517.

Houtekamer PL, Mitchell HL (1998), Data assimilation using an ensemble Kalman filter technique. Mon Weather Rev 126:796–811

Isaaks, E (1990), The application of Monte Carlo methods to the analysis of spatially correlated data. PhD thesis, Stanford University

Jacquard, P., and C. Jain (1965), Permeability distribution from field pressure data, SPEJ 5 (4), 281-294, SPE-1307-PA.

Jafarpour, B., and D. McLaughlin (2009a), Reservoir characterization with discrete cosine transform. part-1: parameterization; part-2 history matching, Soc. Pet. Eng. J., 14(1), 181–202.

Jafarpour, B., and D. McLaughlin, (2009b), Estimating channelized reservoir permeabilities with the ensemble Kalman filter: the importance of the ensemble design. Soc. Pet. Eng. J., 14(2), 374–388.

Jahns, H. O. (1966), A rapid method for obtaining a two-dimensional reservoir description from well pressure response data, Soc. Petrol. Eng. J., 237, 315–27

Jazwinski A (1970), Stochastic processes and filtering theory. Academic Press, San Deigo

Journel, A. (1983), Non-parametric estimation of spatial distributions, Math. Geol. 15(3), 445–468

Journel, A. (2002), Combining knowledge from diverse sources: an alternative to traditional data independence hypotheses, Math. Geol., 34(5), 573–596.

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

Kashib, T., Srinivasan, S., (2006), A probabilistic approach to integrating dynamic data in reservoir models. Journal of Petroleum Science and Engineering 50 (3), 241-257.

Kaviani, D., (2009), Interwell Connectivity Evaluation From Well rate Fluctuations: A Water flooding Management Tool .Ph.D. Dissertation. Texas A & M University, College Station, TX.

Kaviani, D., Jensen, J.L., Lake, L.W. (2012), Estimation of interwell connectivity in the case of unmeasured fluctuating bottomhole pressures, J Pete Sci and Eng 90-91, pp 79-95

LaVenue, A., Ramarao, B., Demarsily, G., and Marietta, M. (1995), Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields. 2. application, Water Resources Research, 31(3):495–516.

Lawson, C. L., and J. R. Hanson (1995), Solving least squares problems, Classics Appl. Math., vol. 15, 337 pp., Soc. for Ind. and Appl. Math., Philadelphia, Pa.

Linden, J., B. Vinsonneau, and K. J. Burnham (2005), Review and enhancement of cautious parameter estimation for model based control: A specific realization of regularization, paper presented at International Conference on Systems Engineering, Inst. of Electr. and Electron. Eng., Las Vegas, Nev., 16– 18 Aug.

Lorentzen, R., G. Nævdal, and A. Lage (2003), Tuning of parameters in a two-phase flow model using an ensemble Kalman filter, International Journal of Multiphase Flow, 29, 1283–1309.

Mariethoz, G., Renard, P., Caers, J., (2010), Bayesian inverse problem and optimization with iterative spatial resampling. Water Resources Research 46 (11).

McLaughlin, D., and L. Townley (1996), A reassessment of the groundwater inverse problem, Water Resour. Res., 32(5), 1131–1161.

Moore, C., and Doherty, J. (2005), Role of the calibration process in reducing model prediction error. Water Resources Research 41, DOI: 10.129/2004WR003501.

Nævdal, G., L.M. Johnsen, S.I. Aanonsen, and E.H. Vefring (2005), Reservoir monitoring and continuous model updating using Ensemble Kalman Filter, SPEJ, 10(1), 66-74.

Nowak, W. (2009), Best unbiased ensemble linearization and the quasi-linear Kalman ensemble generator, Water Resources Research, 45(W04431).

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

Panda, M.N., Chopra,A.K. (1998), An integrated approach to estimate well interactions. Paper SPE 39563 Presented at the SPE India Oil and Gas Conference and Exhibition, New Delhi, India, 17–19 February 1998.

Peaceman, D.W., and Rachford, H.H. (1955), The numerical solution of parabolic and elliptic differential equations. Journal of the Society of Independent Applied Mathematics 3,(1), 28–41.

RamaRao, B., LaVenue, A., de Marsily, G., and Marietta, M. (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.

Reichle R, McLaughlin D, Entekhabi D (2002), Hydrologic data assimilation with the ensemble Kalman filter. Mon Weather Rev 130:103–114

Remy, N., A. Boucher, and J. Wu (2009), Applied geostatistics with SGeMS: A user's guide, Cambridge University Press.

Renard, P. and de Marsily, G.(1997), Calculating equivalent permeability: A review, Advances in Water Resources 20, 253–278.

Renard, P., and J. Caers (2008), Conditioning facies simulations with connectivity data, in Proceedings of the Eighth International Geostatistics Congress, vol. 2, edited by J. M. Ortiz and X. Emery, Gecamin Publ., Santiago, Chile, pp. 597–606.

Reynolds, A., N. He, L. Chu, and D.S. Oliver (1996), Reparameterization techniques for generating reservoir descriptions conditioned to variograms and well-test pressure data. Soc. Petrol. Eng. J., 1(4), 413–426.

Romeu, R. K. and Noetinger, B. (1995), Calculation of internodal transmissibilities in finite difference models of flow in heterogeneous media, Water Resources Research 31(4), 943–959.

Sahuquillo, A., Capilla, J., Gomez-Hernandez, J., and Andreu, J. (1992), Conditional simulation of transmissivity fields honoring piezometric data, In: Hydraulic Engineering Software IV, Fluid Flow Modeling, Kluwer Academic Publishers.

Sambridge, M. and Mosegaard, K. (2002), Monte Carlo methods in geophysical inverse problems, Reviews of Geophysics, 40(3):1009.

Shah P.C., Gavalas G.R., Seinfeld J.H. (1978), Error Analysis in history matching: The optimal level of parameterization, Soc Pet Eng J., 18 (3), 219-228.

Skjervheim, J.A., G. Evensen, S.I. Aanonsen, B.O. Ruud, and T.A. Johansen (2007), Incorporating 4D seismic data in reservoir simulation models using Ensemble Kalman Filter, SPEJ, 12(3), 282-292.

Soeriawinata, T., Kelkar, M. (1999), Reservoir management using production data. Paper SPE 52224 Presented at the SPE Mid-Continent Operations Symposium, Oklahoma City, Oklahoma, 28–31 March 1999.

Solomatine, D. P., and D. L. Shrestha (2009), A novel method to estimate model uncertainty using machine learning techniques, Water Resour. Res., 45, W00B11, doi:10.1029/2008WR006839.

Srivastava, R. M. (1992), Reservoir characterization with probability field simulation, in SPE Annual Technical Conference and Exhibition, Washington, D.C.

Stoyan D, W.S. Kendall, and J. Mecke (1987), Stochastic geometry and its applications, Wiley, New York

Strebelle, S. (2002), Conditional simulation of complex geological structures using multiple-point statistics, Mathematical Geology, 34(1), 1-21.

Sun, N.Z. (1994), Inverse problems in groundwater modeling, Theory and Applications of Transport in Porous Media, 352 pp., Kluwer Academic Publishers, Netherlands.

Tarantola A. (2004), Inverse problem theory. Methods for model parameter estimation, Philadelphia, PA: SIAM.

Thyer, M., B. Renard, D. Kavetski, G. Kuczera, S. W. Franks, and S. Srikanthan (2009), Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis, Water Resour. Res., 45, W00B14, doi:10.1029/2008WR006825.

Tikhonov, A. and V. Arsenin (1977), Solution of ill-posed problems, Winston, PA.

Tonkin, M., and J. Doherty (2005), A hybrid regularized inversion methodology for highly parameterized environmental models, Water Resour. Res., 41, W10412, doi:10.1029/2005WR003995.

Tonkin, M., and J. Doherty (2009), Calibration-constrained Monte Carlo analysis of highly parameterized models using subspace techniques, Water Resour. Res., 45, W00B10, doi:10.1029/2007WR006678.

Weiss, R., and L. Smith (1998), Parameter space methods in joint parameter estimation for groundwater flow models, Water Resour. Res., 34(4), 647– 661, doi:10.1029/97WR03467.

Wen, X., and Chen W. (2006), Real-time reservoir model updating using ensemble Kalman filter, SPEJ, 11 (4), 431–442.

Wen, X. H. and Gomez-Hernandez, J. J. (1996), Upscaling hydraulic conductivities in heterogeneous media: An overview, Journal of Hydrology 183, ix–xxxii.

Western, A.W., Bloschl, G., and Grayson, R.B. (2001), Toward capturing hydrologically significant connectivity in spatial patterns, Water Resour. Res., 37(1), 83–97.

Yeh T.C.J., Lee C.H., Hsu K.C., Illman W.A., Barrash W., Cai X., Daniels J., Sudicky E., Wan L., Li G., Winter C.L., (2008), A view toward the future of subsurface characterization: CAT scanning groundwater basins, Water Resour. Res., 44, W03301, doi:10.1029/2007WR006375.

Yeh, T., C. Lee, K. Hsu, and K. Tan (2007), Fusion of active and passive hydrologic and geophysical tomographic surveys: The future of subsurface characterization, Subsurface Hydrology: Data Integration for Properties and Processes, Geophys. Monogr. Ser., 171, edited by D. W. Hyndman et al.:109.

Yeh, W. (1986), Review of parameter estimation procedures in groundwater hydrology: The inverse problem, Water Resour. Res., 22:95–108.

Yeh, W.W.-G., and Yoon, Y.S. (1981), Aquifer parameter identification with optimum dimension in parameterization. Water Resources Research 17(3), 664–672.

Yousef, A.A, Jensen, J.L., Lake, L.W. (2009), Integrated interpretation of interwell connectivity using injection and production fluctuations. Math. Geosci.41(1), 81–102.

Yousef, A.A., Gentil, P., Jensen, J.L., Lake, L.W. (2006), A capacitance model to infer interwell connectivity from production and injection rate fluctuations. SPEREE 9 (5),630–646.(SPE-95322-PA).

Zhang T, Switzer P, Journel A (2006), Filter-based classification of training image patterns for spatial simulation. Math Geol 38(1):63–80

Zhang, Z., T. Wagener, P. Reed, and R. Bhushan (2008), Reducing uncertainty in predictions in ungauged basins by combining hydrologic indices regionalization and multiobjective optimization, Water Resource. Res., 44, W00B04, doi:10.1029/2008WR006833.

Zhou, H., G_omez-Herna_ndez, J., Li, L., (2012), A pattern-search-based inverse method. Water Resources Research 48 (3), W03505.

Zinn, B., and Harvey, C.F. (2003), When good statistical models of aquifer heterogeneity go bad: A comparison of flow, dispersion, and mass transfer in

connected and multivariate Gaussian hydraulic conductivity fields, Water Resour. Res., 39(3), 1051, doi:10.1029/2001WR001146.