Estimating Reservoir Permeabilities with the Ensemble Kalman Filter: The Importance of Ensemble Design

Behnam Jafarpour, Dennis B. McLaughlin

Department of Civil and Environmental Engineering
Massachusetts Institute of Technology
Cambridge, MA

Corresponding Author
Behnam Jafarpour
Massachusetts Institute of Technology
Department of Civil and Environmental Engineering
Cambridge, MA 02141, USA
behnam@mit.edu

Tel: +1 617-253-1691
Fax: +1 617-258-8850
Summary

Efficient management of smart oilfields requires a reservoir model that can provide reliable forecasts of future production as well as realistic measures of prediction uncertainty. Reliable forecasts depend on an accurate representation of reservoir geology, which is conveyed largely by the permeabilities used in the reservoir simulator. Since these permeabilities cannot be measured directly they must be inferred from measurements of related variables, using procedures such as history matching or Bayesian estimation. The ensemble Kalman filter is an attractive option for permeability estimation in real-time reservoir control applications. It is easy to implement, provides considerable flexibility for describing geological heterogeneity, and supplies valuable information about prediction uncertainty. In this paper we investigate the procedure used to generate the random replicates of an ensemble Kalman filter that estimates reservoir permeabilities. We consider two synthetic water flooding problems based on “true” permeability distributions characterized by conductive channels. The permeability ensembles are obtained from either a classical covariance-based random field generator or a multi-point geostatistical generator. If the ensemble replicates are derived from a covariance model that does not provide for high permeability channels or if the multi-point geostatistical training images do not properly describe the channel geometry the Kalman filter has difficulty identifying the correct permeability field. In fact, in both cases the permeability estimates tend to diverge from the true values as more measurements are included. However, if the filter ensemble replicates are generated by a training image that captures the dominant features of the true permeability field, the filter’s estimates are much better. These results emphasize the importance of generating realistic permeability replicates when using ensemble methods to estimate reservoir properties. In fact, a realistic permeability ensemble appears to be essential for successful estimation.
performance. In practical applications where the true permeability distribution is highly uncertain the prior information used for ensemble generation should properly reflect the full range of possible geological conditions.

Introduction

Parameter estimation is one of the fundamental challenges encountered in real-world reservoir simulation and control applications. The estimation process is greatly complicated by the nature of geological heterogeneity, which is not amenable to simple parametric descriptions. The geological features that govern the flow of liquids in a petroleum reservoir often form preferred pathways that are geometrically complex and difficult to identify from available data. The spatial configuration and properties of these features must be inferred from geophysical, pressure, and flow measurements that are only indirectly related to the parameters of interest. If the independent parameters are assigned to every cell (or pixel) of the simulator’s computational grid the estimation procedure has sufficient flexibility to reproduce more or less arbitrary geometries but the estimation problem may be poorly posed, since many different parameter sets can give comparable fits to available measurements [1-3]. On the other hand, if the number of independent parameters is constrained in an effort to force the problem to be well-posed it may be difficult to adequately describe the true reservoir geometry. Methods for describing the spatial variability of reservoir properties must be both geologically realistic and efficient if the associated parameter estimation procedure is to give satisfactory results. This is true whether the procedure is based on manual adjustment of model inputs or on an automated optimization or history matching procedure.

The parameter estimation approach adopted in a particular application must consider the purpose of the reservoir simulation. Recently, there has been considerable interest in real-
time reservoir control applications where reservoir simulators are used to guide operational activities such as water flooding [4,5]. In such cases real-time measurements are used to continually update the simulator so that it can provide better predictions of the reservoir's response to well adjustments. This, in turn, enables the control algorithm to make the best possible decisions about present and future well settings. Figure 1 shows the basic components of this real-time approach to reservoir operations.

Real-time control requires that simulator parameters and related states be updated frequently, whenever new information becomes available. Automated estimation methods are usually preferred for such applications because of the relatively short turn-around times, the large amount of data to be processed, and the complexity of the estimation problem. In fact, automated real-time parameter and state estimation (or data assimilation) algorithms are now used routinely in many geophysical applications, most notably in weather forecasting [6,7]. Sequential estimation methods such as Kalman filtering [8] are particularly convenient for real-time control problems since they work only with the most recent measurements and estimates, rather than the entire measurement history. However, such methods make assumptions and impose computational demands that may be problematic in some applications.

One of the most popular sequential estimation techniques is the ensemble Kalman filter [9], which is especially well suited for nonlinear problems where model and measurement uncertainties enter in complex ways. In recent years, this method has been introduced to the petroleum engineering literature as a way to estimate uncertain reservoir properties and states [10-12]. The ensemble Kalman filter does not require an adjoint model, is straightforward to implement with commercial reservoir simulators, and is readily parallelized.
Ensemble Kalman filters propagate and update many independent realizations of the uncertain reservoir inputs and states. In non-compositional reservoir simulation applications the model states typically include pressure and saturation. Each realization of the state vector is generated by running the simulator with a particular set of inputs (e.g. permeabilities and other formation properties) drawn at random from specified populations. The input populations should be geologically plausible while including enough variability to properly account for uncertainty. At measurement times input and propagated state replicates can be updated to account for new information. The update process is based on sample covariances derived from the ensemble.

Brouwer et al. [4] used an ensemble Kalman filter to estimate pixel-scale permeabilities from bottom hole pressure observations in a synthetic water flooding experiment. The permeability realizations were generated by randomly sampling a population of exponentially correlated isotropic Gauss-Markov fields with a specified mean, variance, and correlation scale. The true permeability field used to generate the synthetic bottom hole pressure measurements included a preferential pathway that was qualitatively different from any of the sample realizations provided to the filter.

The well controls (pressures and flow rates) used in the Brouwer et al. example were derived from a nonlinear programming algorithm that maximized the net present value of recovered oil, evaluated from the current time to the end of the water flood. The controls varied over time, in response to changing reservoir conditions. This dynamic well excitation tends to make the problem better posed by increasing the sensitivity of the measured states to the uncertain parameters.

The permeability estimates obtained from the ensemble Kalman filter in [4] were able to capture some aspects of the true permeability field at early time but they soon degraded.
In particular, connected channels that were partially identified at early times gradually broke up into disconnected pixels of different permeability values. In this revealing experiment the estimation error actually increased as more information was added. This is a counterintuitive result that could not happen if the system were linear and the estimator optimal. The authors speculated that measurement noise began to dominate the update process as the system approached steady-state conditions behind the advancing water-oil interface. An optimal estimator would handle this situation by decreasing its measurement weights over time.

The ensemble Kalman filter’s ability to properly characterize uncertain reservoir properties depends greatly on the nature of the ensemble it uses. The individual replicates of this ensemble must be geologically realistic and variations across the ensemble must properly capture the dominant sources of uncertainty. If these requirements are not met, the filter’s measurement updates may not be helpful and, in fact, may even be counterproductive. In this paper we reexamine the performance of the ensemble Kalman filter for the water flooding problem. In particular, we consider the benefits of using multi-point geostatistics to generate unconditional permeability ensembles from geologically realistic training images [13,14]. We test this ensemble generation approach on two variants of the Brouwer et al. synthetic water flooding experiment. Our first experiment uses the same “true” permeability field, the same pixel-based parameterization, and the same type of ensemble Kalman filtering algorithm as Brouwer et al. [4]. Our second experiment also relies on a pixel-based parameterization and an ensemble Kalman filtering algorithm but it uses the “true” permeability field adopted by Sarma et al. [5] in their test of a nonlinear least-squares estimation procedure. These two synthetic experiments suggest that the ensemble Kalman filter is able to provide reasonable estimates of geologically realistic permeability fields if the underlying ensemble is properly chosen.
The Ensemble Kalman Filter

We begin with a brief review of the ensemble Kalman filter, as applied to real-time permeability estimation. The Kalman filter can be viewed as a Bayesian estimator that approximates conditional probability densities of the time-dependent state vector $x_t$ [8]. The sequential formulation of the filter distinguishes a forecast (or prior) density $p[x_t | y_{0:t-1}]$ conditioned on all measurements $y_{0:t-1}$ taken prior to time $t$ and an updated (or posterior) density $p[x_t | y_{0:t}]$ conditioned on all measurements $y_{0:t}$ taken through $t$. In our application the state vector consists of the pressure and saturation at the nodes of the spatially discretized simulator computational grid (since capillary pressure is neglected the water and oil pressures are the same). In addition, uncertain model parameters are included in the state vector. This so-called state augmentation approach enables the parameters to be estimated together with the other system states. The measurement vector in our application consists of flow rates and bottom hole pressures measurements inside the wells.

The ensemble Kalman filter approximates the forecast and updated densities with relatively small ensembles of $N$ random realizations, denoted by $x_{t|t-1}^j$ and $x_{t|t}^j$, respectively, where $j = 1, \ldots, N$ represents a particular replicate. The sequence of forecasts and updates is initialized with an ensemble $x_{0|0}^j$ drawn at random from a specified population of initial states. Subsequent forecasts are obtained from the simulator, which may be written as:

$$x_{t|t-1}^j = f_t(x_{t-1|t-1}^j, u_{t-1}, w_{t-1}^j) ; \quad j = 1, \ldots, N$$

(1)
where $u_{t-1}$ is a vector of known (non-random) time-dependent boundary conditions and controls and $w^j_{t-1}$ is a random vector that accounts for uncertain model errors. The function $f_t(\cdot, \cdot, \cdot)$ represents the reservoir simulator, which generates states at $t$ from states and inputs at $t-1$. Time dependent states such as pressure and saturation will generally change over the forecast period while time-invariant states, such as permeabilities, will not.

The updated replicates at $t$ are obtained from a version of the classical Kalman filter update [6,8]:

$$
x^j_{t|t} = x^j_{t|t-1} + \sum_{j=1}^{N} \frac{\text{Cov}[x^j_{t|t-1}, y^j_{t|t-1}] \text{Cov}^{-1}[y^j_{t|t-1}, y^j_{t|t-1}][y_t - y^j_{t|t-1}]}{\text{Cov}[y^j_{t|t-1}, y^j_{t|t-1}]}; \quad j = 1, \ldots, N 
$$

(2)

Here the notation $\text{Cov}[\cdot, \cdot]$ represents the sample covariance between the ensembles associated with the two arguments, $y_t$ is the actual measurement at $t$, and $y^j_{t|t-1}$ represents a perturbed measurement prediction obtained from the following measurement equation:

$$
y^j_{t|t-1} = g_t(x^j_{t|t-1}) + v^j_t; \quad j = 1, \ldots, N
$$

(3)

where $v^j_t$ is a vector of measurement errors, drawn at random from a specified population, and $g_t(\cdot, \cdot)$ relates the measurements at $t$ to the states at $t$. Equations (1), (2), and (3), together with the initial state $x^j_{0|0}$, define the ensemble Kalman filter recursion for the problem of interest here.
The ensemble Kalman filter approach to coupled state/parameter estimation has several characteristics which deserve to be mentioned. First, the method offers the flexibility of generating the random realizations $x_{0|0}^j$, $w_t^j$, and $v_t^j$ from any desired population. In the real-time reservoir control context, this means that permeability realizations included in $x_{0|0}^j$ may be drawn from a population of physically realistic alternatives. These populations can, for example, be constructed from libraries of preferred channel configurations that share certain distinctive features.

Another attractive characteristic of the ensemble Kalman filter is its ability to generate non-Gaussian sample distributions of the states $x_{t|t-1}^j$ and $x_{t|t}^j$. There is no need to linearize or otherwise approximate the state transition function $f_t(\cdot, \cdot, \cdot)$ or to assume that the random states and inputs are Gaussian. However, the ensemble Kalman filter converges to the true conditional densities $p[x_t | y_{0:t-1}]$ and $p[x_t | y_{0:t}]$, only when all prior states and measurements are jointly Gaussian. This typically only occurs when the state and measurement equations are both linear. In most petroleum reservoir applications the joint Gaussian condition is not met and the filter’s sample densities and moments are only approximations. In practice, the updated sample mean $E[x_{t|t}^j]$ is typically used as a point estimate for characterization and control purposes. This is the estimate that we will be examining when comparing different ensemble generation approaches.

It should be noted that the covariance inversion operation in (2) is computationally expensive and can be ill-conditioned for large problems. This can be avoided by using the pseudo inverse procedure based on singular value decomposition, as proposed in [6]. Also, several variants of the EnKF have been developed to improve the computational
efficiency of the algorithm and to address some of the issues in its implementation, mostly due to sampling errors (see [6 -7] for a brief review).

Experimental Setup

We revisit the water flooding example originally studied in [4] and later examined in [5]. In this example a $450(m) \times 450(m) \times 10(m)$ synthetic reservoir is discretized into a two-dimensional $45\times45\times1$ uniform grid block system, as shown in Figure 2. The simulations are performed with the commercially available ECLIPSE [15] reservoir simulator, which is set up for two phase (oil and water) black oil flow. The total simulation time is 1080 days, divided into 12 intervals of 90 days. Horizontal wells with 45 ports are used to inject water into the left side of the reservoir and to produce oil and water from the right side end.

The injection wells are operated with specified flow rates while the production wells are operated with specified bottom hole pressures. In [4] and [5] the well port settings at each simulation time were determined by an optimization algorithm designed to maximize the net present value of benefits obtained from oil recovery. Here we specify the port settings beforehand and focus on permeability estimation rather than optimal recovery.

A total of one pore volume of water is injected into the reservoir during the simulation. The injection and production wells are each divided into 3 different groups of 15 well ports and the simulation is divided into 6 time periods, with each period lasting 180 days. The ports are represented in Figure 2 by small colored circles on either side of the simulation domain. Each column of circles shows the well settings used in one of the 6 operating periods. The total amount of injected water is divided evenly over the 6 production periods and among the injection ports open in any given period. The injection strategy can be summarized as follows:
Injection ports:

- Periods A (0-180 days) and E (720-900 days) – Uniform distribution of 1/6 PV of water among all ports.
- Periods B (180-360 days), and D (540-720 days) – Uniform distribution of 1/6 PV of water among Groups 1 and 3 respectively.
- Periods C (360-540 days), and F (900-1080 days) – Uniform distribution of 1/6 PV of water among Group 2.

The production strategy is:

Production ports:

- Periods A and E – All production ports have a specified bottom-hole pressure of 2990 psi.
- Periods B and C – Production ports in Groups 1 and 3 have a specified bottom-hole pressure of 2990 psi while the pressure at other ports is kept at 3000 psi.
- Periods D and F – Production ports in Group 2 have a specified bottom-hole of 2990 psi while pressure at other ports was kept at 3000 psi.

In this study the only source of simulator uncertainty is the permeability, which is treated as a random field. Initial and boundary conditions are assumed to be known perfectly and dynamic model errors are assumed to be negligible. In situations where these assumptions may not hold additional error sources may be included in the ensemble filtering process. The initial reservoir pressure and connate water saturation are 3000 psi and 0.10, respectively, throughout the reservoir. The realizations used to construct the permeability ensemble are generated with the *snesim* algorithm of the Stanford Geostatistical Modeling Software (S-GeMS) [14]. This software relies on a multiple-point
geostatistical method that uses a specified training image to determine the general structure of the realizations. In our experiments different training images were used to evaluate the impact of the ensemble on filter performance.

The ensemble Kalman filter uses two types of measurements in its updates of state and input replicates: 1) bottom hole pressure observations at each of the 45 ports in the injection wells and 2) oil and water flow rate measurements at each of the 45 ports in the production wells. In each experiment the “true” injection well bottom hole pressures and production well flow rates are generated by running the simulator from a specified “true” permeability field. Uncorrelated zero mean random measurement errors are added to these “true” pressures and flow rates. The standard deviations of the random measurement errors are 20 psi and 20 sbpd for bottom hole pressures and flow rates, respectively.

The ensemble filter estimates the natural log of permeability, which is then transformed to permeability for input to the simulator. Jafarpour et al. [16] discuss some of the limitations of log permeability transformations for parameter estimation applications. We use this transformation here to maintain consistency with [4] and [5]. In some cases the classical ensemble Kalman filter update can give unphysical saturation values outside the range [0, 1]. In order to avoid this problem, our Kalman filter works with the transformed saturation \( S^* \), which is distributed over \(( -\infty, +\infty)\) and is computed as follows:

\[
S^* = erf^{-1}(2 \cdot S - 1) \quad \rightarrow \quad S = \frac{1}{2} erf(S^* + 1)
\]

where \( erf \) represents the error function. After the filter update \( S^* \) is transformed back to the saturation \( S \) for use in the reservoir simulator.


Experiments and Discussion

Experiment 1.

This experiment uses the “true” log permeability field adopted in the Brouwer et al. synthetic water flooding study [4] and shown in Figure 3a. This “true” field includes a distinctive high permeability channel that cuts through the reservoir from the injection to production sides (see [4] for more details). The log permeability ensemble generated by Brouwer et al. [4] consists of 100 realizations of a Markov random field with a Gaussian spatial correlation function. The correlation length is a normally distributed random variable with a mean of 20 grid blocks and a standard deviation of one grid block. The mean and standard deviation of the log normally distributed random permeability are 200 mD and 1.5 mD, respectively. The “true” field is qualitatively different in structure than the log permeability ensemble since it contains a channel that does not appear in any of the ensemble replicates.

In our version of the Brouwer et al. synthetic water flooding experiment we use two different permeability ensembles generated with multi-point geostatistics. These are based on two different training images, each consisting of two facies -- sandstone and shale. The sandstone and shale permeabilities are 10,000 mD and 500 mD, respectively. These are comparable in magnitude to the channel and mean background permeability in the Brouwer et al. field. The two log permeability training images and some typical replicates are shown in Figures 3b and 3c, respectively. Note that the training images cover a significantly larger region than the simulation domain.

Ensemble Kalman filters rely on a limited number of samples (or replicates) drawn from a specified population or probability distribution. If the sample size is too small the resulting
sample statistics may be inaccurate and the filter’s performance may suffer [17,18]. The dependence of estimation accuracy on ensemble size is application-specific. In our study we performed a sensitivity analysis to identify an ensemble size that yields robust permeability estimates with the least possible computational cost. This study indicated that the filter results for our synthetic experiments converge only if the ensemble size is at least 300. Based on these results we selected an ensemble size of 300 for our study. This is significantly larger than the 100 replicate ensemble used in [4].

In Experiment 1-1 we used Training Image 1 (Figure 3b) to generate the ensemble members for estimation. This training image has wider channels than the true log permeability field used in [4]. It is the same image used in the synthetic water flooding experiment described in Sarma et al. [5]. Figure 4b shows the log permeability estimate (first row) and log permeability estimation error (the difference between “true” and ensemble mean log permeability, second row) obtained after each analysis step with an ensemble derived from Training Image 1. As seen in these figures, the ensemble estimate fails to capture the spatial structure of the high permeability channel. Also, the initial updates are better than the later ones, as was observed by Brouwer et al. [4].

Figure 4b also shows the standard deviation of the updated log permeability ensemble (third row) and the ensemble mean of the oil saturation (fourth row). The standard deviation is generally highest in the middle of the domain where the estimation pixels are furthest from the measurements. The filter’s ensemble standard deviation significantly underestimates the actual estimation error, indicating that the filter is overconfident. This overconfidence tends to make the filter ignore information from the measurements. The mean oil saturation differs most from the true saturation in the beginning of the simulation, where it misses the water front advancing through the high permeability channel. Near the
end of the simulation the estimate and the true saturation values tend to converge, despite the relatively poor quality of the permeability estimate. This appears to reflect the fact that the water front moves slower at the two ends of the reservoir, where the permeability is underestimated while it moves faster in the middle where the permeability is overestimated. Overall, the average front speed is reasonably close to the true case.

The log permeability ensemble used in Experiment 1-1 is based on a training image characterized by high permeability channels that are consistently wider than the channel appearing in the “true” field. The average channel width in the true field is about 3-4 grid blocks (30-40 meters) while the Training Image 1 channels have an average channel width of about 10 grid blocks (100 meters). As a result, channels in the ensemble members generated from the Training Image 1 are consistently too wide and the sample covariances used to update these replicates do not adequately describe the spatial features of the log permeability field. It is difficult to predict the effects of such ensemble specification errors. In the particular example considered here our synthetic simulation indicates that these errors are sufficient to significantly degrade filter performance.

The importance of the training image and the resulting log permeability ensemble can be demonstrated if we run a new Experiment 1-2, which is the same as Experiment 1-1 except that we use Training Image 2 (Figure 3c). Training Image 2 has narrower channels and gives a better description of the true channel geometry. The results of Experiment 1-2 are shown in Figure 5. It is readily apparent that the filter is better able to capture the channel. However, the log permeability error plots indicate that the position of the estimated channel is displaced somewhat in the middle of the domain. This likely reflects the fact that different geometries in this region provide nearly the same fits to the measurements, which are taken at the edges of the simulation domain. So the filter has
difficulty converging to the “true” channel position, even though it correctly infers that there is a channel in the general vicinity. The ensemble standard deviation is noticeably smaller in Experiment 1-2 than in Experiment 1-1. This indicates that the replicates are less dispersed around the mean in Experiment 1-2.

Experiments 1-1 and 1-2 suggest that the gradual degradation in the permeability estimates obtained by Brouwer et al. in [4] can be explained by ensemble selection. The replicates used to derive sample covariances in [4] and in our Experiment 1-1 do not adequately capture dominant features in the “true” log permeability field. This leads to incorrect updates that can cause the filter estimates to drift away from the true values, even as more measurements are added. When the replicates provide a better characterization of reality, as in our Experiment 1-2, the ensemble Kalman filter works much better.

Another contributing factor that could have caused the reported divergence of the EnKF in [4] may be associated with the size of the ensemble used. The sensitivity analysis in our study indicated that the ensemble size of 100 used in [4] does not provide a sufficiently accurate representation of the prior covariance used to derive updates. This further aggravates the covariances errors introduced by an inappropriate ensemble specification. Although larger ensemble sizes imply more computational effort, it may be argued that the increased effort is justified by the improved results. Also, the efficiency of the ensemble Kalman filter can be significantly improved, especially for larger problems, if parallel computation and more advanced solution procedures are adopted.

The improvement in the updated permeability enhances the predictive power of the reservoir model. This, in turn, can be used to improve decisions regarding well location or
production strategies. Past studies have reported improved oil recovery when permeability estimation and optimal control approaches are combined [4,5]. These studies have concluded that an approximate estimate of the permeability field (especially proper identification of low and high permeability features) gives significant improvement in the sweep efficiency. It is noteworthy that the degradation in log permeability estimates observed in Brouwer et al. [4] and in our Experiment 1-1 occurred at later update times after the most important control decisions had already been made. It would be useful to investigate to what extent these decisions will change when a more accurate estimate of the log permeability field is used to predict recovery.

**Experiment 2**

In order to further investigate the connection between the log permeability ensemble and filter performance we consider another water flooding experiment that uses two “true” permeabilities drawn from an ensemble derived from Training Image 1 (Figure 3b).

Figures 6 and 7 summarize the results obtained for Experiments 2-1 and 2-2, which use the “true” permeability images shown in Figures 6a and 7a, respectively. These “true” images correspond to Realizations 9 and 22 of the Training Image 1 ensemble (this ensemble is drawn from the library used by Sarma et al. [5]). The channel geometries in these cases are somewhat more complex than in the Experiment 1 “true” image but the sandstone channels (10,000 mD) and background shale (500 mD) permeabilities are uniform rather than spatially variable. The estimation results summarized in Figures 6 and 7 confirm that the ensemble Kalman filter is able to identify the general structure of the log permeability field when the “true” permeability and ensemble are compatible. The filter captures most of the general features after the first two updates and there is no
performance degradation with time. In fact, the quality of the estimates tends to improve until the end of the simulation (36 months) when the final solution is obtained. The log permeability estimates produced in Experiment 2-1 are somewhat better than in Experiment 2-2, probably because the image in Experiment 2-2 is more complex, especially in the center region that is further from the measurements. Experiments with other “true” log permeability images (not shown here) confirm that complex features located further from measurements are more difficult to identify.

In both “true” cases there are errors in the position of the estimated channel, similar to the errors observed in Experiment 1-2. These position offsets are evident in both the mean error and standard deviation plots, which mirror the geometry of the “true” channel. Channel position offsets result in high, and sometimes misleading, root-mean square errors since the differences between the “true” and ensemble mean log permeability values in the vicinity of a position offset can be as large as the “true” value in the channel. In this case, the root-mean squared errors on the edges of the channel can actually be higher if the channel is identified, but the estimated location is shifted slightly, than if the channel is missed altogether. The critical question here is “what error measure best reflects the quality of a permeability estimate in a channelized setting?”. This is a topic that deserves further investigation.

Conclusions

The synthetic water flooding experiments conducted in this study support the view that proper definition of the permeability ensemble is essential to the success of the ensemble Kalman filter in reservoir applications. Geological features in real reservoirs have complex channelized geometries that cannot generally be reproduced with statistically
homogeneous Markov random fields. Permeability generation methods based on training images and multi-point geostatistical methods seem to be better able to generate geologically credible realizations in channelized environments. The ensemble Kalman filter uses these realizations to derive sample covariances that provide approximate statistical descriptions of spatial variability. Our experiments indicate that the sample covariances, and the updates that depend on them, give better results when the underlying ensemble members have geometrical properties similar to the true permeability.

Experiment 1-1 shows that it is not enough to use training images that generate channels. The channels that appear in the ensemble replicates must have dimensions and other geometrical properties that are compatible with the true log permeability. In Experiment 1-1 a training image approach produced relatively poor estimates that degraded over time, in a manner similar to the estimates described by Brouwer et al. [4]. The channels in the ensemble replicates for this case were consistently too wide. In Experiment 1-2 the channels in the training image were narrowed to be closer in width to the channel included in the true permeability field. In this case, the prior information conveyed by the ensemble was sufficiently accurate to give significantly better results. The dominant channel present in the “true” field was identified, although the position was slightly shifted.

The ensemble Kalman filter’s ability to identify high permeability channels was confirmed in Experiment 2, where two different synthetic “true” channel configurations were considered. In both cases, the training image was compatible with the “true” image and the estimation results were encouraging. Features nearer the well measurements were generally estimated more accurately than features located further away.

The performance of the EnKF is also dependent upon the number of realizations used in the ensemble. In our example, a log permeability ensemble size of 100 was too small to
give reliable results while an ensemble size of 300 seemed to be sufficient. The sample covariances that control the filter updating procedure can be expected to improve when the ensemble replicates are realistic and when the ensemble is large enough to provide an adequate characterization of uncertainty.

It is reasonable to ask whether we can expect an ensemble generated from a specified training image to include replicates that look like the unknown “true” log permeability field in a real application. Ensemble generation is definitely more challenging in a real application than it was in our Experiment 2, where the true permeability is one of the replicates in the training image library. In more realistic situations it is important that the training image (or images) used for permeability estimation be derived from field data at the site of interest and that the image be sufficiently large, complex, and diverse to include all the features likely to occur at the site. At locations where the geology is highly uncertain the training image channels should vary in width, tortuosity, connectedness, and complexity, so that this uncertainty is reflected in the ensemble. Proper ensemble design is a critical part of the parameter estimation process. If the filter’s ensemble reflects the likely range of true conditions the resulting estimates can be expected to be more accurate and robust. This is an important topic that deserves further investigation from both research and application perspectives.
Nomenclature

\( \text{Cov}[\ldots] = \) Sample covariance matrix.
\( E[\cdot] = \) Expectation operator.
\( \text{erf}(\cdot) = \) Error function operator.
\( f(\ldots) = \) State propagation function (multiphase flow simulation equation).
\( g(\cdot) = \) Measurement equation.
\( j = \) Superscript used for replicate indexing.
\( N = \) Ensemble size (number of replicates used in an ensemble).
\( p[\cdot] = \) Probability density.
\( p[\cdot | \cdot] = \) Conditional probability density.
\( S = \) Saturation of a known phase.
\( S' = \) Transformed saturation.
\( t = \) Subscript indicating time step.
\( u = \) Vector of known (non-random) time dependent inputs.
\( v = \) Vector of time dependent measurement errors.
\( w = \) Random vector representing time dependent uncertain model errors.
\( x = \) Vector containing all states in the model.
\( y = \) Actual measurement vector.
\( | = \) Probability conditioning operator.

Acknowledgments

This research has been funded by Shell International Exploration and Production. Schlumberger's ECLISPE E100 reservoir simulator was used for multiphase flow simulation.
References


**SI Metric Conversion Factors**

\[
\begin{align*}
\text{bbl} & \times 1.589873 \times 10^{-1} = \text{m}^3 \\
\text{ft} & \times 3.048 \times 10^{-1} = \text{m} \\
\text{psi} & \times 6.894757 \times 10^0 = \text{kPa}
\end{align*}
\]
Author Biographies

Behnam Jafarpour received a B.S. degree from Tehran University, Iran-1999, and a M.S. degree from the University of Delaware, USA-2003, both in Civil and Environmental Engineering. He is currently a Ph.D. candidate in Civil and Environmental Engineering and a M.S. student in Electrical Engineering at the Massachusetts Institute of Technology (MIT), Cambridge. His research interests include model parameterization, data assimilation, and control theory with application to reservoir engineering.
E-mail: behnam@mit.edu

Dennis B. McLaughlin received a Ph.D. in Civil and Environmental Engineering from Princeton University, Princeton, NJ, in 1985. He is currently the H. M. King Bhumibol Professor of Water Resources Management, Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT), Cambridge. His research interests include environmental data assimilation and management of water resources in semi-arid regions. He is a Fellow of the American Geophysical Union.
E-mail: dennism@mit.edu
Figure Captions

Figure 1
Reservoir closed-loop control body diagram adapted from [4]. Model updating (history matching) component is shown in the bottom right loop.

Figure 2
Reservoir configuration with injection/production schemes in a waterflooding experiment. Injection wells with dark blue, light blue, and cyan circles have high, intermediate, and low injection rates, respectively. The red and yellow circles in the production wells indicate low and high bottom hole pressures, respectively.

Figure 3
(a) True log-permeability for Experiment 1; (b) Log permeability training images used in Experiments 1 and 2. Training Image 1 has wide channels and is used in Experiments 1-1, 2-1, and 2-2. It is shown in the first row along with nine sample permeability realizations, generated by SGeMS using multiple point geostatistics. Training Image 2 has narrower channels and is used in Experiment 1-2. It is shown in the second row along with nine sample permeability realizations generated by SGeMS.

Figure 4
(a) True log-permeability field in Experiment 1 (first row) and its corresponding saturation field (second row) for the injection/production scenarios shown in Figure 2; (b) Estimated mean (first row), error (second row), and standard deviation (third row) of the log-permeability field and estimated mean saturation field (fourth row) for Experiment 1-1 after
the EnKF updates at 0, 3, 6, 18, and 36 months. The initial permeabilities were generated from Training Image 1 with wide channels (Fig. 3b)

**Figure 5**
(a) True log-permeability field in Experiment 1 (first row) and its corresponding saturation field (second row) for the injection/production scenarios shown in Figure 2; (b) Estimated mean (first row), error (second row), and standard deviation (third row) of the log-permeability field and estimated mean saturation field (fourth row) for Experiment 1-2 after the EnKF updates at 0, 3, 6, 18, and 36 months. The initial permeabilities were generated from Training Image 2 with narrow channels (Fig. 3b)

**Figure 6**
(a) True log-permeability field in Experiment 2-1 (first row) and its corresponding saturation field (second row) for the injection/production scenarios shown in Figure 2; (b) The estimation mean (first row), error (second row), and standard deviation (third row) for the log-permeability field and the estimated mean saturation field (fourth row) for Experiment 2-1 after the EnKF updates at 0, 3, 6, 18, and 36 months. The initial permeabilities were generated from Training Image 1 with wide channels (Fig. 3b)

**Figure 7**
(a) True log-permeability field in Experiment 2-2 (first row) and its corresponding saturation field (second row) for the injection/production scenarios shown in Figure 2; (b) The estimation mean (first row), error (second row), and standard deviation (third row) for the log-permeability field and the estimated mean saturation field (fourth row) for Experiment 2-2 after the EnKF updates at 0, 3, 6, 18, and 36 months. The initial permeabilities were generated from Training Image 1 with wide channels (Fig. 3b).
Figure 1

Adapted from Brouwer et al. [4]
Figure 2

Reservoir Configuration

Injection and Production Schedules:

A: 0-180 days          D: 540-720 days
B: 180-360 days        E: 720-900 days
C: 360-540 days        F: 900-1080 days
Figure 3

a) Experiment 1 “True” Log Permeability

b) Log Permeability Training Images Used in Experiments 1 and 2
Figure 4

a) Experiment 1 “True” Values

b) Experiment 1-1 Estimation Results, Training Image 1
Figure 5

a) Experiment 1 “True” Values

b) Experiment 1-2 Estimation Results, Training Image 2
Figure 6
a) Experiment 2-1 “True” Values

b) Experiment 2-1 Estimation Results, Training Image 1
Figure 7

a) Experiment 2-2 “True” Values

b) Experiment 2-2 Estimation Results, Training Image 1