An Iterative Ensemble Kalman Filter for Multiphase Fluid Flow Data Assimilation

Yaqing Gu, SPE, BP; and **Dean S. Oliver,** SPE, U. of Oklahoma

Summary

The dynamical equations for multiphase flow in porous media are highly non-linear and the number of variables required to characterize the medium is usually large, often two or more variables per simulator gridblock. Neither the extended Kalman filter nor the ensemble Kalman filter is suitable for assimilating data or for characterizing uncertainty for this type of problem. Although the ensemble Kalman filter handles the nonlinear dynamics correctly during the forecast step, it sometimes fails badly in the analysis (or updating) of saturations.

This paper focuses on the use of an iterative ensemble Kalman filter for data assimilation in nonlinear problems, especially of the type related to multiphase flow in porous media. Two issues are key: (1) iteration to enforce constraints and (2) ensuring that the resulting ensemble is representative of the conditional pdf (i.e. that the uncertainty quantification is correct). The new algorithm is compared to the ensemble Kalman filter on several highly nonlinear example problems, and shown to be superior in the prediction of uncertainty.

Introduction

For linear problems, the Kalman filter is optimal for assimilating measurements to continuously update the estimate of state variables. Kalman filters have occasionally been applied to the problem of estimating values of petroleum reservoir variables (Eisenmann et al. 1994; Corser et al. 2000), but they are most appropriate when the problems are characterized by a small number of variables and when the variables to be estimated are linearly related to the observations. Most data assimilation problems in petroleum reservoir engineering are highly nonlinear and are characterized by many variables, often two or more variables per simulator gridblock.

The problem of weather forecasting is in many respects similar to the problem of predicting future petroleum reservoir performance. The economic impact of inaccurate predictions is substantial in both cases, as is the difficulty of assimilating very large data sets and updating very large numerical models. One method that has been recently developed for assimilating data in weather forecasting is ensemble Kalman filtering (Evensen 1994; Houtekamer and Mitchell 1998; Anderson and Anderson 1999; Hamill et al. 2000; Houtekamer and Mitchell 2001; Evensen 2003). It has been demonstrated to be useful for weather prediction over the North Atlantic. The method is now beginning to be applied for data assimilation in groundwater hydrology (Reichle et al. 2002; Chen and Zhang 2006) and in petroleum engineering (Nævdal et al. 2002, 2005; Gu and Oliver 2005; Liu and Oliver 2005a; Wen and Chen 2006, 2007; Zafari and Reynolds 2007; Gao et al. 2006; Lorentzen et al. 2005; Skjervheim et al. 2007; Dong et al. 2006), but the applications to state variables whose density functions are bimodal has proved problematic (Gu and Oliver 2006).

For applications to nonlinear assimilation problems, it is useful to think of the ensemble Kalman filter as a least squares method that obtains an averaged gradient for minimization not from a variational approach but from an empirical correlation between model variables (Anderson 2003; Zafari et al. 2006). In addition to providing a mean estimate of the variables, a Monte Carlo estimate of uncertainty can be obtained directly from the variability in the ensemble.

Assume that at time t_k, we have an ensemble of samples of state and model vectors from the posterior $p(f_k, m_k|D_{\text{obs,k}})$ where $D_{\text{obs,k}} = \{D_{\text{obs,k}-1}, d_{\text{obs,k}}\}$ is the collection of all data through time t_k. Bayes theorem relates the probability density for the state variables, f_k, and model variables, m_k, after assimilation of data $d_{\text{obs,k}}$ at time t_k to the prior probability density at t_k as follows:

$$\begin{aligned} p(f_k,m_k|D_{\text{obs,k}}) &\propto p(d_{\text{obs,k}}|f_k,m_k)p(f_k,m_k|D_{\text{obs,k}-1}) \\ &\propto p(d_{\text{obs,k}}|f_k,m_k)p(f_k|m_k,D_{\text{obs,k}-1}) \\ &\quad \times p(m_k|D_{\text{obs,k}-1}) \\ &\propto p(d_{\text{obs,k}}|m_k)p(m_k|D_{\text{obs,k}-1}) \quad \ldots . \quad (1) \end{aligned}$$

The equivalence is a result of the fact that the state variables f_k can be computed from the model variables m_k. Although the first line of Eq. 1 is fundamental for the traditional Kalman filter because it explains how to update the model and state variables directly, the third line simply points out the possibility of using the Kalman filter to update only the model variables (and initial conditions), then computing the state from the model variables. It seems that the uncertainty in initial conditions is relatively small in most petroleum reservoir engineering applications because the reservoir is typically in a state of static equilibrium at the time of first production. Exceptions include uncertainty in the initial fluid contacts, uncertainty in irreducible water saturation (Tjølsen et al. 1994), and the possibility of tilted initial fluid contacts. If important, these types of uncertainty can be treated by including additional variables in the state vector (Evensen et al. 2007; Thulin et al. 2007).

The posterior pdf from which samples should be drawn (Eq. 1) is the product of two terms. The prior is represented by the ensemble of model and state vectors. As in the traditional EnKF, for purposes of updating the variables, we approximate the prior by a Gaussian whose mean and covariance are estimated from the ensemble. The other term is the likelihood; when both terms are Gaussian, the product is Gaussian and the Kalman filter can be used to compute the updated model and state variables. When the likelihood is not Gaussian (e.g. when the relationship between model variables and observation variables is nonlinear), the product is not a Gaussian, but sampling from an approximation to the posterior can still be accomplished fairly efficiently.

For a problem in which the relationship between the state variables, the model parameters, and the data is linear, both the model parameters and the state variables can be updated simultaneously using the Kalman filter. The result is an improved estimate of the (non-varying) model parameters and also an improved estimate of the current value of the state variables.

For a nonlinear problem, it may be impossible to update the state variables to be consistent with the updated model parameters without re-solving the nonlinear forward problem to obtain state variables. In many applications of the ensemble Kalman filter, the objective is primarily to estimate the current state of the system. For petroleum reservoir applications, however, it is generally important to estimate not only the current state of the system (the pressures and saturations), but also the values of permeability, porosity, and fault transmissibility.

Model variables, m, are variables that are uncertain but are not time varying. These include rock properties such as absolute permeability and porosity. The estimates of these properties change as data are assimilated, but the parameter itself should not be interpreted to be changing with time. If initial conditions and boundary conditions are uncertain, they can be included in the list of model variables. State variables, f, are uncertain, time-dependent variables that define the state of the system. For petroleum reservoirs, these could include pressure in each fluid phase, saturations

Original SPE manuscript received for review 11 May 2007. Revised manuscript received 13 June 2007. Paper (SPE 108438) peer approved 9 July 2007.

of phases, or mass fraction. The state variables are frequently solutions of systems of differential (or difference) equations. If the model is valid, and the initial conditions and model variables are known, then it is possible to compute the state variables for any time. Data, d, are observable quantities related to the state of the model and indirectly to the model parameters. For petroleum reservoirs, data might include bottomhole pressure (possibly at several locations in the well-bore), surface flow rates, and amplitude of seismic reflection. Theoretical data can be computed from the state and model variables. Observations always have some level of measurement error or noise associated with them.

Application of the ensemble Kalman filter to petroleum reservoir flow problems is much simplified by the introduction of a joint model-state-observation vector consisting of model variables, state variables, and theoretical data (Tarantola 1987; Anderson 2001). In a typical application of the ensemble Kalman filter to a two-phase petroleum reservoir flow and transport problem, we might define an augmented state vector, Y, of the form,

$$Y = \left[\phi^{\mathrm{T}}, \ln k^{\mathrm{T}}, P^{\mathrm{T}}, S_w^{\mathrm{T}}, WOR^{\mathrm{T}}\right]^{\mathrm{T}}, \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2)$$

which has been partitioned into model variables that do not change with time (porosity, ϕ, and log-permeability, $\ln k$), state variables that change substantially (pressure, P, and water saturation, S_w), and predictions of observations or theoretical data (producing water-oil ratio, WOR). In our history matching (and some geophysical inverse theory) terminology, we denote the static model variables by the symbol m, the dynamic state variables at time t_k by $f_k(m)$, and the predictions of observations at time t_k by $g_k(m)$. In this nomenclature, the augmented state vector at time t_k is

$$Y_k = \begin{bmatrix} m \\ f_k(m) \\ g_k(m) \end{bmatrix}. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (3)$$

Because the primary focus of this paper is on the assimilation step, we will omit the time subscript, understanding that the functions $f(\cdot)$ and $g(\cdot)$ are generally functions of time. The relationship between the observations and the true static model variables is

$$d_{\mathrm{obs}} = g(m_{\mathrm{true}}) + \varepsilon \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4)$$

where ε is the measurement error is assumed to be Gaussian and $E[\varepsilon\varepsilon^{\mathrm{T}}] = C_D$. The relationship between the observations and the true augmented state vector can also be written as

$$d_{\mathrm{obs}} = HY_{\mathrm{true}} + \varepsilon, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (5)$$

where H can be represented as a matrix whose elements are all ones or zeroes.

Linear Dynamic System. If the relationships between the model variables, the state variables, and the theoretical data are linear, then the augmented state vector can be written as

$$Y = \begin{bmatrix} m \\ Fm \\ Gm \end{bmatrix} = \begin{bmatrix} I \\ F \\ G \end{bmatrix} m \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

where F and G are linear operators. If the model variables m are multivariate Gaussian with auto-covariance C_M, then the auto-covariance of the augmented state vector is

$$\begin{aligned} C_Y &= \begin{bmatrix} I \\ F \\ G \end{bmatrix} C_M \begin{bmatrix} I^{\mathrm{T}} & F^{\mathrm{T}} & G^{\mathrm{T}} \end{bmatrix} \\ &= \begin{bmatrix} C_M & C_M F^{\mathrm{T}} & C_M G^{\mathrm{T}} \\ FC_M & FC_M F^{\mathrm{T}} & FC_M G^{\mathrm{T}} \\ GC_M & GC_M F^{\mathrm{T}} & GC_M G^{\mathrm{T}} \end{bmatrix}. \quad \ldots\ldots\ldots\ldots \quad (7) \end{aligned}$$

After acquisition of new data d_{obs}, the estimate of the model variables with minimum variance is

$$\langle m \rangle = m_{\mathrm{pr}} + C_M G^{\mathrm{T}} (GC_M G^{\mathrm{T}} + C_D)^{-1} (d_{\mathrm{obs}} - Gm_{\mathrm{pr}}). \quad \ldots \quad (8)$$

This solution can also be written in terms of the augmented state vector

$$\langle Y \rangle = Y_{\mathrm{pr}} + C_Y H^{\mathrm{T}} (HC_Y H^{\mathrm{T}} + C_D)^{-1} (d_{\mathrm{obs}} - HY_{\mathrm{pr}}) \quad \ldots. \quad (9)$$

because $HC_Y H^{\mathrm{T}} = GC_M G^{\mathrm{T}}$. Eq. 9 is the analysis or updating step in the Kalman filter. It simultaneously updates the model variables, the state variables, and the estimate of the data. Approximations to the products $HC_Y H^{\mathrm{T}}$ and $C_Y H^{\mathrm{T}}$ can be efficiently computed from the ensemble of state vectors (Evensen 2003).

Nonlinear Dynamic System. For nonlinear dynamic system, we can consider a series of linear approximations to the nonlinear functions, $f(m)$ and $g(m)$, by linearizing them at point m^ℓ. Suppose that $g(m) \approx g(m^\ell) + G_\ell (m - m^\ell)$ and $f(m) \approx f(m^\ell) + F_\ell (m - m^\ell)$, so the approximation to the covariance based on linearization at m^ℓ is

$$C_Y \approx \begin{bmatrix} C_M & C_M F_\ell^{\mathrm{T}} & C_M G_\ell^{\mathrm{T}} \\ F_\ell C_M & F_\ell C_M F_\ell^{\mathrm{T}} & F_\ell C_M G_\ell^{\mathrm{T}} \\ G_\ell C_M & G_\ell C_M F_\ell^{\mathrm{T}} & G_\ell C_M G_\ell^{\mathrm{T}} \end{bmatrix}. \quad \ldots\ldots\ldots\ldots \quad (10)$$

The vector of model variables that maximizes the conditional probability density also minimizes the following objective function.

$$\begin{aligned} S(m) &= \frac{1}{2}\left(g(m) - d\right)^{\mathrm{T}} C_D^{-1} \left(g(m) - d\right) \\ &\quad + \frac{1}{2}\left(m - m_{\mathrm{pr}}\right)^{\mathrm{T}} C_M^{-1} \left(m - m_{\mathrm{pr}}\right). \quad \ldots\ldots \quad (11) \end{aligned}$$

In this notation, m_{pr} denotes the estimate of m at the end of the forecast step (before the assimilation or analysis step). When the number of data is smaller than the number of model variables, the most appropriate form of the Gauss-Newton method for finding the $\ell + 1$ iterative estimate of the model vector that minimizes the objective function in Eq. 11 is

$$\begin{aligned} m^{\ell+1} &= m_{\mathrm{pr}} - C_{\mathrm{M}} G_\ell^{\mathrm{T}} (C_{\mathrm{D}} + G_\ell C_{\mathrm{M}} G_\ell^{\mathrm{T}})^{-1} \\ &\quad \times \left[g(m^\ell) - d_{\mathrm{obs}} - G_\ell (m^\ell - m_{\mathrm{pr}})\right]. \quad \ldots \quad (12) \end{aligned}$$

When the number of data is larger than the number of model variables, the most appropriate form is

$$\begin{aligned} m^{\ell+1} &= m_{\mathrm{pr}} - (G_\ell^{\mathrm{T}} C_{\mathrm{D}}^{-1} G_\ell + C_{\mathrm{M}}^{-1})^{-1} G_\ell^{\mathrm{T}} C_{\mathrm{D}}^{-1} \\ &\quad \times \left[g(m^\ell) - d_{\mathrm{obs}} - G_\ell (m^\ell - m_{\mathrm{pr}})\right]. \quad \ldots \quad (13) \end{aligned}$$

On the other hand, if the problem is sufficiently nonlinear that a reduced step length is required, the Gauss-Newton formula for iteration is

$$\begin{aligned} m^{\ell+1} &= \beta_\ell m_{\mathrm{pr}} + (1 - \beta_\ell) m^\ell - \beta_\ell C_{\mathrm{M}} G_\ell^{\mathrm{T}} (C_{\mathrm{D}} + G_\ell C_{\mathrm{M}} G_\ell^{\mathrm{T}})^{-1} \\ &\quad \times \left[g(m^\ell) - d_{\mathrm{obs}} - G_\ell (m^\ell - m_{\mathrm{pr}})\right] \quad \ldots. \quad (14) \end{aligned}$$

where β_ℓ is an adjustment to the step length whose optimal value can be determined by standard methods (Dennis and Schnabel 1983). These formulas have been the basis for most of the Gauss-Newton or Levenberg-Marquardt approaches for automatic history matching (Gavalas et al. 1976; Tan and Kalogerakis 1992; Li et al. 2003).

Note that C_M in Eqs. 12 and 14 is the model covariance prior to assimilation of the current data but after assimilation of all data before the current time. It does not change during the Gauss-Newton iteration, although the linear approximation to the measurement operator, G, may change with each iteration. As a result, the computation of the product $G_\ell C_{\mathrm{M}} G_\ell^{\mathrm{T}}$ is not as straightforward as in the EnKF where it is only necessary to compute $(H\Delta Y)(H\Delta Y)^{\mathrm{T}}$ (Evensen 2003).

Iterative forms of the Kalman filter are not entirely new. The iterated Kalman filter (Jazwinski 1970) is in fact similar in character to what we proposed here, Ensemble Randomized Maximum Likelihood Filter (EnRML), although without the ensemble. The standard Kalman filter attempts to estimate the conditional mean, but for highly nonlinear problems, the conditional mode is often a

more appropriate measure of the central tendency. Bell and Cathey (1993) discussed the equivalency of the iterated Kalman filter to the Gauss-Newton method for approximating the maximum likelihood estimate. Also, Zupanski (2005) discusses a maximum likelihood ensemble filter (MLEF) that is viewed as a maximum likelihood approach to the ensemble transform Kalman filter of Bishop et al. (2001). In the MLEF method, the maximum likelihood model and the Hessian are estimated.

Several ad hoc iterative applications of the ensemble Kalman filter have been proposed in the petroleum engineering literature. Gu and Oliver (2006) used iteration to reduce the magnitude of the nonlinear effects on the saturation correction. Wen and Chen (2007) proposed a conforming EnKF method in which the flow simulator was rerun from a previous step to generate plausible saturation values. Liu and Oliver (2005b) iterated on the Kalman correction to enforce nonlinear constraints for facies observations with a truncated plurigaussian model. Recently, Zafari et al. (2006) review the EnKF filter through the lens of optimization, to derive an iterative EnKF procedure for nonlinear problems.

This paper includes five test problems designed to validate various aspects of the proposed iterative filter, including verification that it gives correct results on linear problems.

Problem 1 verifies that the results with reduced and full step length are identical to results from the analysis step of the EnKF for a linear problem.

Problem 2 verifies that the estimates of the mean and the variance from the new iterative analysis step are much better than the results from the analysis step in EnKF for a nonlinear problem.

Problem 3 tests the impact of the ensemble size on the estimation of the "sensitivity matrix" from the ensemble for a nonlinear, 10-parameter problem. The ensemble is smaller than the number of model variables in some cases so the estimate of G is underdetermined, as it will be in most real cases.

Problem 4 is a linear dynamic problem in which a passive tracer is injected and the time of breakthrough is measured. Both the dynamics and the observation operator are linear so the results can be compared with EnKF, which should be correct.

Problem 5 is a nonlinear dynamic problem on which the methods can be tested, but for which the correct results are not known. We can, however, show that the state variables (saturation) from the EnRML are physically plausible and match the data.

Implementation of the EnRML. Let $M_{\rm pr}$ be the matrix whose columns consist of the ensemble of model vectors after assimilation of all previous data. There are N_e of these vectors, and hence N_e columns of $M_{\rm pr}$. Denote the vector of means of the prior variables by $\bar{m}_{\rm pr}$ and the matrix of deviations from the means by $\Delta M_{\rm pr}$. The ensemble estimate of the prior model variable covariance (after assimilation of all previous data) is $C_{\rm M} = \Delta M_{\rm pr}\Delta M_{\rm pr}^{\rm T}/(N_e-1)$. One feature that makes the implementation of the traditional ensemble Kalman filter so efficient is that it is never necessary to compute $C_{\rm M}$, only the products $HC_{\rm Y}H^{\rm T}$ and $C_{\rm Y}H^{\rm T}$. This computation is not as straightforward in an iterative filter, because it is important to maintain the distinction between the model covariance matrix estimate, which should be based on the prior models, and the sensitivity matrix, which should be based on the current values. At the ℓth iteration, let ΔD^{ℓ} represent the deviation of each vector of computed data from the mean vector of computed data and let ΔM^{ℓ} represent the deviation of each vector of model variables from the current mean. The ensemble average sensitivity matrix G_{ℓ} is the coefficient matrix relating the changes in model parameters to the changes in computed data,

$$\Delta D^{\ell} = G_{\ell}\Delta M^{\ell} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (15)$$

where ΔM^{ℓ} is $N_M \times N_e$; ΔD^{ℓ} is $N_D \times N_e$; G_{ℓ} is $N_D \times N_M$. N_M is the number of the model parameters and N_D is the number of data. As ΔM^{ℓ} is not generally invertible (or even square), we use the singular value decomposition (SVD) to solve the system. Although the dimension of the model variables is large, the size of the ensemble is typically fairly small so the effort required to compute the SVD is reasonable (Golub and van Loan 1989, p. 239).

Except for the iterative aspects, the EnRML procedure is very similar to the EnKF procedure and we use the first step of the recursive process to illustrate the EnRML procedure in the following.

1. Compute the reservoir state variables using the updated model parameters, $m_{\rm j,0}^{\ell}$, from the initial time 0 to the first measurement time t_1

$$\psi(m_{\rm j,0}^{\ell}, t: 0 \rightarrow t_1) \quad j = 1,2,\ldots,N_e \quad \ldots\ldots\ldots \quad (16)$$

   where $\psi(\cdot)$ denotes the reservoir simulator, the first argument specified in the parentheses are model parameters used to re-initialize the flow equations at initial time 0. The state variables for the re-initialization are determined by the initial conditions. The second subscript on m is the time index.

2. Generate an approximation of the sensitivity matrix by "solving" Eq. 15 for G_{ℓ} using a pseudo-inverse based on singular value decomposition of ΔM^{ℓ}.

3. Apply the Gauss-Newton update formula in Eq. 14 to improve estimates of model variables, $m_{j,0}^{\ell+1}(j = 1,2,\ldots,N_e)$.

4. Evaluate the data mismatch term for both $m_{j,0}^{\ell}$ and $m_{j,0}^{\ell+1}$

$$S(M_0) = \sum_{j=1}^{N_e} \left(g(m_{\rm j,0}) - d_{\rm obs,j,1}\right)^{\rm T} \times C_D^{-1}\left(g(m_{\rm j,0}) - d_{\rm obs,j,1}\right). \quad \ldots. \quad (17)$$

   Note the computation of $g(m_{\rm j,0}^{\ell+1})$ involves solving the forward flow equations again from initial time 0 with $m_{\rm j,0}^{\ell+1}$

$$\psi(m_{\rm j,0}^{\ell+1}, t: 0 \rightarrow t_1) \quad j = 1,2,\ldots,N_e \ldots\ldots\ldots \quad (18)$$

5. If $S(M_0^{\ell+1}) < S(M_0^{\ell})$, overwrite m^{ℓ} with $m^{\ell+1}$ and increase β_{ℓ}; otherwise, keep m^{ℓ} and decrease β_{ℓ}.

6. Check if the convergence criteria are satisfied. If not, go to **Step 1** and iterate the procedure; otherwise, exit the assimilation step and begin the next forecast.

The following criteria are used to determine if the solution has converged:

- $\mathbf{MAX}_{1\le i\le N_M; 1\le j\le N_e} \mid m_{i,j}^{\ell+1} - m_{i,j}^{\ell} \mid < \varepsilon_1$ or
- $S(M^{\ell+1}) - S(M^{\ell}) < \varepsilon_2 S(M^{\ell})$ or
- $S(M^{\ell+1}) \le N_D$ or
- Iteration exceeds the pre-set maximum number of iterations, $I_{\rm MAX}$.

In this paper, $\varepsilon_1 = 10^{-5}$, $\varepsilon_2 = 10^{-4}$, $I_{\rm MAX} = 20$ for the linear problem, and $I_{\rm MAX} = 6$ for the nonlinear problem.

A comprehensive description of the non-iterative EnKF procedure, including refinements that can be used to decrease the effects of small sample size can be found in Evensen (2003).

Test Problems

Because the key feature of the EnRML approach is the iteration in the analysis step, the first several examples do not involve integration of dynamical systems. Subsequent tests apply the method to linear and nonlinear dynamical flow and transport problems.

Problem 1: Single-Variable Linear, Analysis Only. The ensemble-based implementation of the Randomized Maximum Likelihood method of sampling (Kitanidis 1995; Oliver et al. 1996) should perform quite well on a linear problem with large ensembles, because both the variance and the cross-covariance can be estimated accurately.

TABLE 1—SUMMARY STATISTICS PROBLEM 1 (LINEAR)				
Estimate	EnKF	EnRML (0.5)	EnRML (1.0)	True
$\overline{\text{mean}}$	0.000	0.000	0.000	0
$\overline{\text{var}}$	0.498	0.498	0.498	0.5
$\overline{\text{obj}}$	99.97	99.97	99.97	100
$\overline{\text{iterations}}$	1	12.2	2	NA

TABLE 2—SUMMARY STATISTICS FOR PROBLEM 2 (NONLINEAR)				
Estimate	EnKF	EnRML (0.5)	EnRML (1.0)	True
$\overline{\text{mean}}$	-2.04	-2.80	-2.80	-2.84
$\overline{\text{var}}$	0.033	0.069	0.070	0.067
$\overline{\text{iterations}}$	1	22.7	9.8	NA

In this simple linear example, a single measurement is made of a random variable drawn from a standard normal distribution. The goal is to estimate the value of the variable and the uncertainty of the estimate. The observed value is 0, but it is assumed to be contaminated with Gaussian measurement noise with mean 0 and variance 1. To reduce the effect of sampling error, an ensemble of 100 states is generated to test the performance of both the EnKF and the EnRML methods. For this linear test problem, the variable itself is observed so $g(m) = m$, but the observation is noisy, so $d_{\text{obs}} = g(m_{\text{true}}) + \varepsilon$.

Prior to acquisition of data, the best estimate of m is 0, and the uncertainty is completely characterized by the variance, $\sigma_m^2 = 1$. After assimilation of the data, the best estimate is still 0, but the posterior variance is $1/2$. For each ensemble, after the analysis step, the mean, the variance, and the data objective function,

$$\text{obj} = \sum_{j=1}^{100} \frac{(g(m_{\text{i,j}}) - d_{\text{obs}})^2}{C_\varepsilon}$$

are computed for each ensemble. The values $\overline{\text{mean}}$, $\overline{\text{var}}$, and $\overline{\text{obj}}$ reported in **Table 1** are the averages over 10,000 ensembles. Not surprisingly, results from the EnKF method are quite good. The iterated filter also gives the correct results, at a somewhat higher price. In particular, note that the iterated filter (EnRML) with restricted step length ($\beta = 0.5$) took an average of 12 iterations to converge, but obtained results that were identical to those of the EnKF. This verifies that the iterated update, even with a restricted step, gives correct results for a linear problem.

Problem 2: Single-Variable Nonlinear, Analysis Only. In this example, a single measurement is made of a random variable drawn from a standard normal distribution. The observed value is $g(-3)$, but the measurement is assumed to be contaminated with Gaussian noise with mean 0 and variance 0.01. An ensemble of 100 states is generated to test the methods with relatively small sampling error. We tested the analysis step from both the EnKF and the EnRML. Two values of β were tested in the iterative update, simply to verify consistency. For this nonlinear test problem,

$$g(m) = m + (m/3)^2.$$

Prior to acquisition of data, the best estimate is $m^{\text{pr}} = 0$ and the variance in the estimate is $\sigma_m^2 = 1$. In this case, the results for EnRML with the full step length ($\beta = 1$) and for EnRML with restricted step length ($\beta = 0.5$) are nearly identical, except that the restricted step length took twice as many iterations to converge (**Table 2**). Results from the EnRML are much better than results from the EnKF for this problem. In particular, note that the variance estimate from EnRML is very close to the correct result, while the estimate from EnKF is too small by a factor of 2.

Problem 3: Ten Variable Nonlinear, Analysis Only. One important aspect of the iterative updating method is the need to compute (and update) the sensitivity matrix for the data G when the number of model variables is larger than the number of ensemble members.

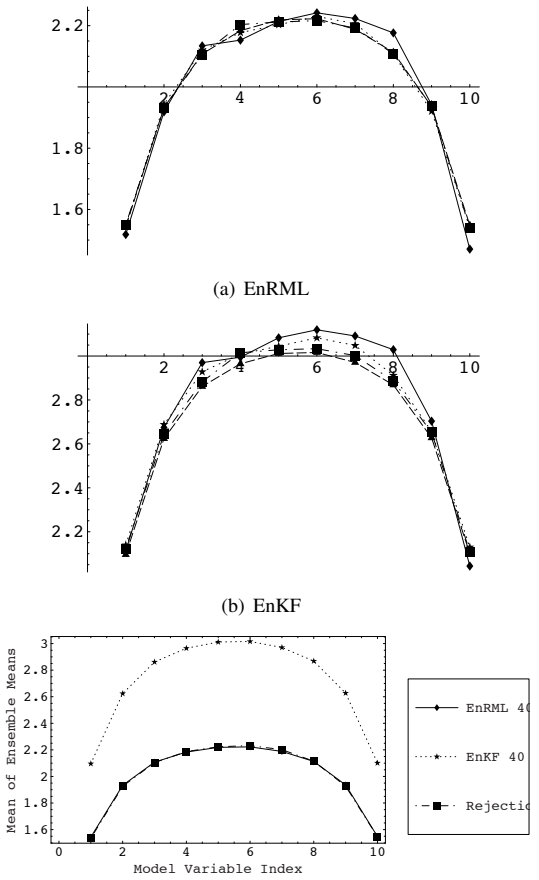


Fig. 1—Mean of ensemble means from 10,000 trials with ensemble size varying from 5 (diamonds), 10 (stars), 20 (squares), 40 (triangles), 80 (dashed).

The test problem has a larger number of variables than the ensemble size, to test the impact of the estimation of G from a small ensemble. The variables form a 1-D gaussian random field on a uniform lattice. The prior expectation for the variables is $m_{\text{pr}} = \{0, \ldots, 0\}$ and the prior covariance is $(C_M)_{i,j} = \exp(-3|i-j|/4)$. (Note that the prior variance is uniform, but that the covariance is not.)

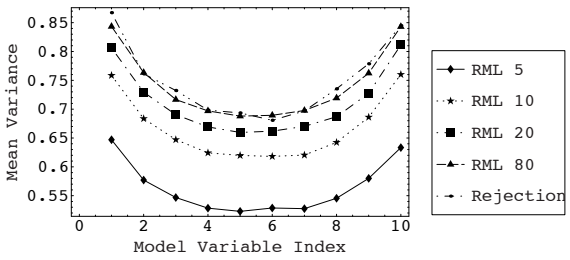
A single nonlinear measurement is made of the quantity

$$g(m) = \bar{m} + 0.2\bar{m}^2$$

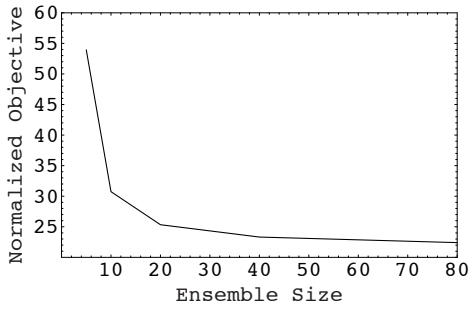
where $\bar{m} = (m_1 + \cdots + m_{10})/10$. We assume that the measurement is $d = g(\{2, \ldots, 2\}) = 2.8$ and that the standard deviation of the measurement error is 0.01. (This implies that the mean of the model variables must be approximately 2.0.)

Means and variances for each of the 10 variables were estimated using standard implementations of the analysis step in EnKF and the EnRML. Results from the EnKF and EnRML were compared with results from an acceptance/rejection algorithm (Ripley 1987) in which 10 million random samples were proposed from a multigaussian approximation to the posterior pdf. 22,000 of the proposals were accepted.

Ensembles for EnKF and EnRML varied in size from 5 to 40 and the experiment was repeated 10,000 times for each ensemble size. **Fig. 1** shows the means of the ensemble estimates for each of the variables from the EnKF and EnRML. Note that the results for the mean do not seem to depend significantly on the size of the ensemble, but results from the EnKF are much different from those of EnRML, which are nearly identical to results from the rejection method (**Fig. 1c**). The estimate of variance from the ensemble after



(a) Mean variance as a function of ensemble size.



(b) Mean normalized objective function as a function of ensemble size.

Fig. 2—Results from 10,000 trials with ensemble size varying from 5 (diamonds), 10 (stars), 20 (squares), 40 (triangles), 80 (dashed).

the EnRML iterative update does depend on the size of the ensemble, but for an ensemble size of 80, the results are nearly identical to the results from the rejection algorithm (**Fig. 2a**). The normalized objective function for an ensemble of realizations m_i with perturbed data $d_{\text{obs,i}}$ is the mean of the following over the 10,000 trials:

$$S(m) = \frac{1}{2N_e}\sum_{i=1}^{N_e}\left[\left(g(m_i)-d_{\text{obs,i}}\right)^2 C_D^{-1} + \left(m_i - m_{\text{pr,i}}\right)^{\text{T}} \times C_M^{-1}\left(m_i - m_{\text{pr,i}}\right)\right]. \quad \ldots\ldots\ldots\ldots \quad (19)$$

The magnitude of the mean normalized objective function initially decreases rapidly as the size of the ensemble increases, but increasing the size above 40 results in little further improvement for this problem (**Fig. 2b**).

Problem 4: Linear Dynamics and Observation Operator. In this problem, dynamics are included so that the result of iterating from time 0 can be compared with the standard EnKF without iteration, and with another iterative filter that does not require a return to the initial time. Assume that a passive tracer is injected at one end of a core and that the times for transport of the tracer to different locations of the core is measured. The flow is single phase; the length of the core is L; the cross-sectional area is A; the permeability is $k(x)$; the porosity is $\phi(x)$; and the viscosity of the fluid is μ. For simplicity, assume also a consistent set of units, so Darcy's Law is $q = uA = kA\Delta p(x)/(\mu x)$, where q is the flow rate, u is the superficial velocity, and $\triangle p(x)$ is the pressure drop from the inlet to location x. The velocity of the tracer front is $v = u/\phi$.

For fixed flow rate q, the arrival time of the tracer at location x is

$$t_x = \frac{x}{v} = \frac{A}{q}\int_0^x \phi(x')dx', \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (20)$$

or for uniformly discretized grids,

$$t_x = \frac{A\triangle x}{q}\sum_{i=1}^{i_x}\phi_i, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (21)$$

where i_x is the discretized grid index corresponding to the location x. We can see from Eq. 21 that the travel time of the tracer to the grid i_x depends on the summation of the porosity of the core from the inlet up to the grid i_x. It does not depend on the porosity elsewhere, or on the permeability or the fluid viscosity.

For this selected problem, the core is discretized into 20 uniform grids. The arrival times of the tracer at the downstream ends of grids 4, 7, 12, and 20 are used to update the porosity distribution, and the tracer concentrations of the core. To propagate the system from previous assimilation time (i.e. previous measured location) is

$$t_{x_k} = t_{x_{k-1}} + \frac{A\triangle x}{q}\sum_{i=1+i_{x_{k-1}}}^{i_{x_k}}\phi_i$$

and to propagate the system from time 0 (location 0) is

$$t_{x_k} = \frac{A\triangle x}{q}\sum_{i=1}^{i_{x_k}}\phi_i,$$

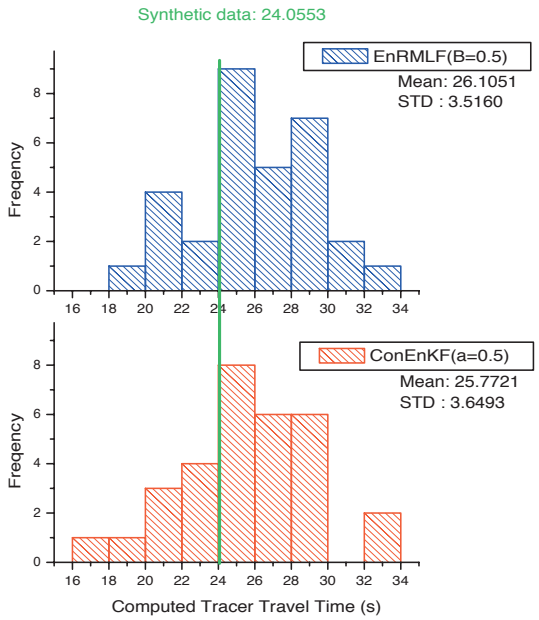
where $i_{x_{k-1}}$ and i_{x_k} are the integer grid indices corresponding to the location x_{k-1} and x_k, respectively.

Results from three methods were compared for this problem: EnKF, EnRML, and conforming EnKF (Wen and Chen 2007). Conforming EnKF is an iterative form of the ensemble Kalman filter in which only the model variables are updated using the Kalman gain. The reservoir simulator is used to update the state variables, by restarting the reservoir simulator from the previous assimilation time with the updated model variables. By using the simulator to update the state variables, the values are physically plausible. The EnRML method is somewhat similar, except that the iterative scheme is developed from the Gauss-Newton formulation of the Randomized Maximum Likelihood method, which has been extensively tested by comparison with Markov chain Monte Carlo results (Liu and Oliver 2003). EnRML does demand greater computation as it requires the state variables to be recomputed from initial conditions. Differences between the methods were accentuated by assuming that the first measurement of arrival time was highly inaccurate ($\sigma_\varepsilon = 100.0$), while subsequent measurements were more accurate ($\sigma_\varepsilon = 0.25$). Histograms of the computed data after assimilation of the first and third observation are shown in **Fig. 3**. There is very little difference between the results from the two methods after the first data assimilation (**Fig. 3a**), but the differences are pronounced (**Fig. 3b**) after the assimilation of the third data. Residuals from EnRML at the third assimilation time are consistent with the assumption of small noise in the observations.

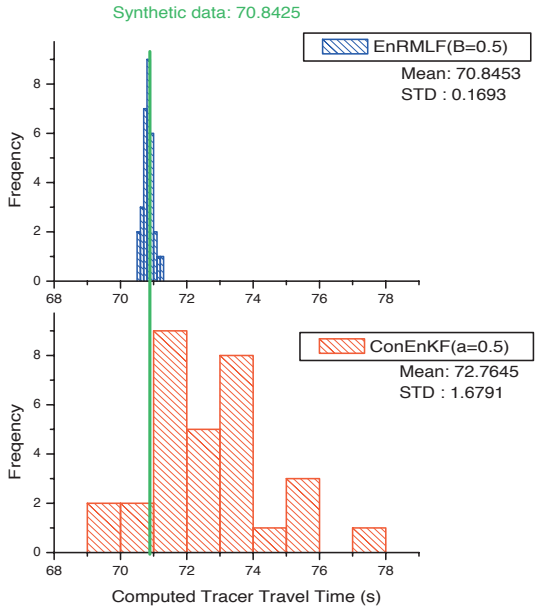
Fig. 4 compares the mean of the ensemble porosity estimates with 30 members from the three methods at measurement time 4. The vertical green line indicates the measurement location. The results from the EnKF and EnRML overlay each other, indicating that for a problem with linear dynamics and a linear observation operator, EnRML gives results that are identical to the results from EnKF.

Problem 5: Nonlinear Dynamics and Observation Operator. This example deals with data assimilation for 1-D, two-phase, immiscible flow without capillary pressure. The problem is chosen because the saturation shock results in a bimodal probability density for saturation that is difficult to handle with the standard Kalman filters. The test case has 32 grid cells in a 1D grid. Water is injected at a constant rate in grid 1 and fluid is produced at a constant pressure from grid 32. Water saturations are measured at the observation well in grid 16.

An exponential covariance model is used to generate 65 initial realizations of porosity and log permeability. The practical correlation range (the distance at which the covariance drops to 5% of the variance value (Journel and Huijbregts 1978) of the covariance model is approximately 15 grids. The mean and standard deviation of porosity fields are 0.2 and 0.04, respectively. The mean and standard deviation of $\ln k$ fields are 5.5 and 0.7. The correlation coefficient between porosity and $\ln k$ is 0.6.



(a) At measurement time 1



(b) At measurement time 3

Fig. 3—Histograms of the computed tracer travel times with 30 ensemble members from the Conforming EnKF and EnRML both with half step length at the end of their iterations at measurement times 1 and 3.

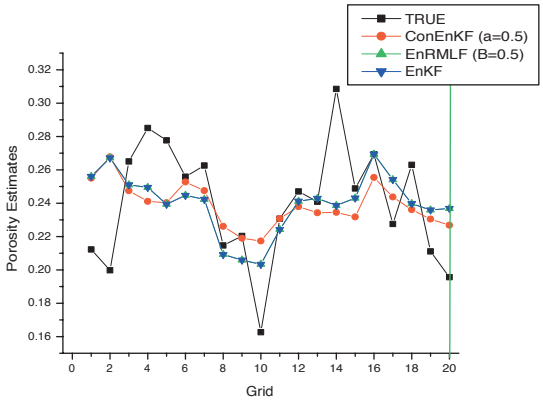


Fig. 4—Ensemble mean porosity with 30 ensemble members from the EnKF, conforming EnKF and EnRML both with half step length at the end of their iterations at measurement time 4.

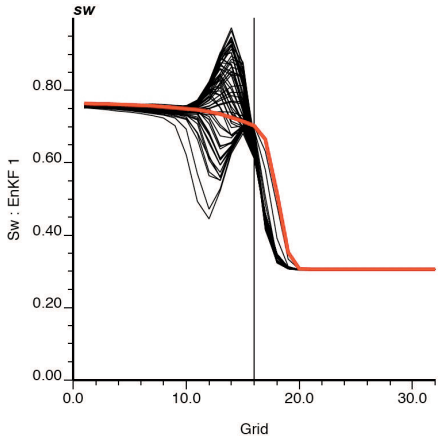
In a previous investigation, it was found that the ensemble Kalman filter worked fairly well at updating the water saturation field when corrections to the saturation were small, but that the updated values became non-physical when the measurements started at a late time (Gu and Oliver 2006). For this test, water saturation observations at days 40, 90, 140, 190, 210, 230, 250, and 270 are used to refine the reservoir models. None of the models have water breakthrough at the measurement location before 190 days, so there are no changes to the model or state variables until 190 days. At that time, the correction is fairly large. The standard deviation of the error contained in the data is 0.5%, a small value that was chosen to make the problem more nonlinear.

Fig. 5 shows the ensemble of water saturation profiles after the assimilation of data at day 190. The vertical line indicates the observation location (grid 16). The red curve is the true saturation profile at this time. The multiple black lines are saturation profiles from the ensemble. The EnKF gives saturation profiles that are completely implausible; in some cases water saturations exceed $1 - S_{or}$ and in others the saturation profiles oscillate between high and low values (**Fig. 5a**). Results from the conforming EnKF are plausible but, even after six iterations, inconsistent with the observed saturation (**Fig. 5b**). Results from the EnRML are both plausible and consistent with the observed saturation at the measurement location (**Fig. 5c**).

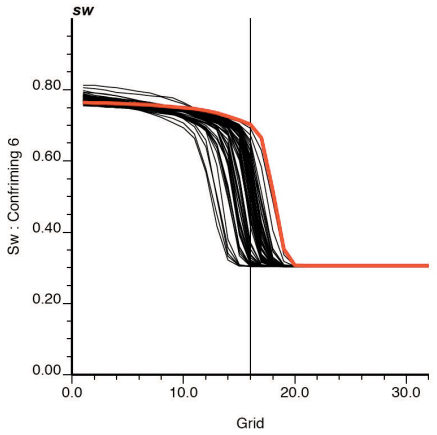
In each of these methods, both the model variables (porosity and log-permeability) and the state variables (saturation and pressure) are updated to be consistent with observations. The saturation profiles in **Fig. 5** show, however, that the errors in saturation in the neighborhood of the front can be quite large for both EnKF and conforming EnKF. The errors in the porosity are generally smaller and more consistent between the methods. We compare the estimates of porosity quantitatively using the root-mean-square error (RMSE) in the porosity variables.

$$\mathrm{RMSE}_k = \sqrt{\frac{1}{N_e}\sum_{j=1}^{N_e}(x_{k,j} - x_k^{\mathrm{true}})^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (22)$$

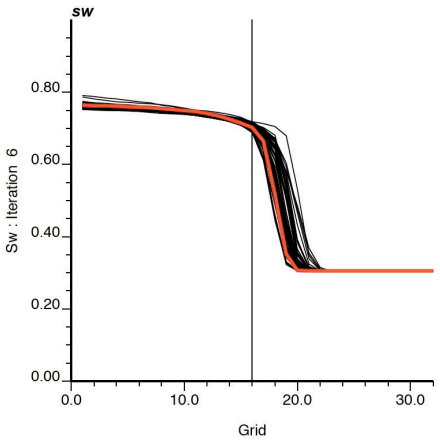
Fig. 6 plots the individual gridblock RMSE in the porosity values at 190 days (**Fig. 6a**) and the maximum value of the gridblock RMSE at each of the measurement times (**Fig. 6b**). Although there is no particular reason that the RMSE in model parameters should always decrease as data are assimilated, we can make a couple of general observations. Despite the large differences in saturation estimates between EnKF and EnRML, the errors in the porosity estimates are very similar. None of the methods make any changes in the model until 190 days, at which time all of the methods show a substantial



(a) After EnKF correction.



(b) After 6 conforming EnKF iterations.



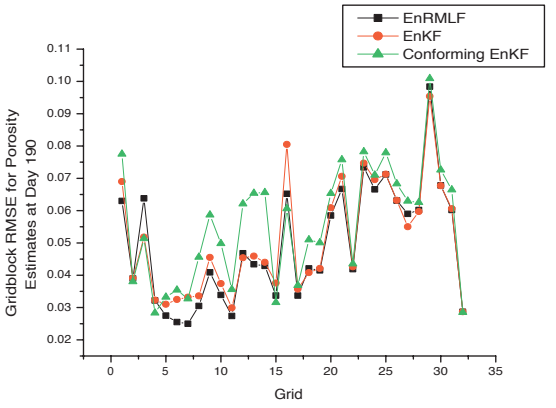
(c) After 6 EnRML iterations.

Fig. 5—Assimilation of S_w observation at 190 days.

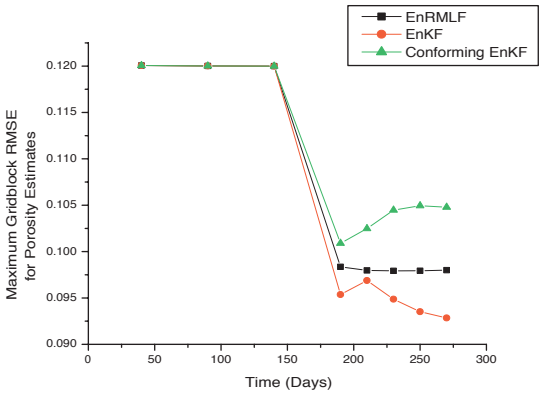
reduction in the RMS error. In this example, the maximum RMSE in porosity after 200 days is smallest for the EnKF.

Discussion and Conclusions

Although the EnKF has generally proven to be effective for history matching and assimilation of data in multiphase fluid flow problems, there are times when the resulting state estimates or realizations are inconsistent with the data, or are inconsistent with the dynamical equations and physical limits. The solutions that we have



(a) RMSE in porosity at 190 days.



(b) Maximum RMSE in porosity at all assimilation times.

Fig. 6—Comparison of porosity values for EnKF, conforming EnKF, and EnRML.

proposed resolve the issue of inconsistency in the state variables by generating the updated (history matched) state variables from the dynamical equations instead of from the filter — only the static variables are updated from the filter. Because of the nonlinearity in the relationship between model and state variables, and between model variables (e.g. porosity and permeability) and observable variables (e.g. water saturation, pressure, rates), it is necessary to iterate to enforce the measurement constraints. We demonstrated through a number of examples that the iterative filter (EnRML) gave the correct results for small linear static problems, and agreed with the EnKF for a linear dynamical problem involving the flow of tracer in a porous medium.

We also compared the results on two static nonlinear problems, and showed that the estimates of the mean and the variance from EnRML were very close to the correct values, while results from EnKF were not. For the dynamic problem of two-phase flow in a porous medium, with observations of saturation, we showed that EnRML, EnKF, and conforming EnKF gave similar results for porosity (and permeability) but that the estimate of the state variable (saturation) was much better from EnRML.

The new EnRML method is more expensive than EnKF because of the requirement to rerun the simulation models from the beginning time to update the estimates of saturation and pressure. In practice (Gu 2006), we have found that it is best to use the EnKF correction when the corrections to the state variables are small, and to use the EnRML correction when the changes are large. In this

way, estimates and realizations of the state variables and the model variables are consistent with all of the data and the physical constraints.

References

Anderson, J. L. 2001. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Monthly Weather Review* **129** (12): 2884–2903. DOI

Anderson, J. L. 2003. A Local Least Squares Framework for Ensemble Filtering. *Monthly Weather Review* **131** (4): 634–642.

Anderson, J. L. and Anderson, S. L. 1999. A Monte Carlo Implementation of the Nonlinear Filtering Problem to Produce Ensemble Assimilations and Forecasts. *Monthly Weather Review* **127** (12): 2741–2758.

Bell, B. M. and Cathey, F. W. 1993. The iterated Kalman Filter Update as a Gauss-Newton Method. *IEEE Transactions on Automatic Control* **38** (2): 294–297.

Bishop, C. H., Etherton, B. J., and Majumdar, S. J. 2001. Adaptive Sampling With the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Mon. Wea. Rev.* **129**: 420–436.

Chen, Y. and Zhang, D. 2006. Data Assimilation for Transient Flow in Geologic Formations via Ensemble Kalman Filter. *Advances in Water Resources* **29** (8): 1107–1122.

Corser, G. P., Harmse, J. E., Corser, B. A., Weiss, M. W., and Whitflow, G. L. 2000. Field Test Results for a Real-Time Intelligent Drilling Monitor. Paper SPE 59227 presented at the IADC/SPE Drilling Conference, New Orleans, 23–25 February. DOI: 10.2118/59227-MS.

Dennis, J. E. and Schnabel, R. B. 1983. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs.

Dong, Y., Gu, Y., and Oliver, D. S. 2006. Sequential Assimilation of 4D Seismic Data for Reservoir Description Using the Ensemble Kalman Filter. *Journal of Petroleum Science and Engineering* **53** (1–2): 83–99.

Eisenmann, P., Gounot, M.-T., Juchereau, B., and Whittaker, S. J., 1994. Improved Rxo Measurements Through Semi-Active Focusing. Paper SPE 28437 presented at the SPE Annual Technical Conference and Exhibition, New Orleans, 25–28 September. DOI: 10.2118/28437-MS.

Evensen, G. 1994. Sequential Data Assimilation With a Nonlinear Quasi-geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics. *Journal of Geophysical Research* **99** (C5): 10143–10162.

Evensen, G. 2003. The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation. *Ocean Dynamics* **53**: 343–367.

Evensen, G., Hove, J., Meisingset, H. C., Reiso, E., Seim, K. S., and Espelid, O. 2007. Using the EnKF for Assisted History Matching of a North Sea Reservoir Model. Paper SPE 106184 presented at the SPE Reservoir Simulation Symposium, Houston, 26–28 February. DOI: 10.2118/106184-MS.

Gao, G., Zafari, M., and Reynolds, A. C. 2006. Quantifying Uncertainty for the PUNQ-S3 Problem in a Bayesian Setting with RML and EnKF. *SPEJ* **11** (4): 506–515. SPE-93324-PA. DOI: 10.2118/93324-PA.

Gavalas, G. R., Shah, P. C., and Seinfeld, J. H. 1976. Reservoir History Matching by Bayesian Estimation. *SPEJ* **16** (6): 337–350. SPE-5740-PA. DOI: 10.2118/5740-PA.

Golub, G. H. and van Loan, C. F. 1989. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, second edition.

Gu, Y. 2006. *History Matching Production Data Using the Ensemble Kalman Filter*. PhD dissertation, University of Oklahoma.

Gu, Y. and Oliver, D. S. 2005. History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter. *SPEJ* **10** (2): 51–65. SPE-89942-PA. DOI: 10.2118/89942-PA.

Gu, Y. and Oliver, D. S. 2006. The Ensemble Kalman Filter for Continuous Updating of Reservoir Simulation Models. *Journal of Energy Resources Technology* **128** (1): 79–87.

Hamill, T. M., Snyder, C., Baumhefner, D. P., Toth, Z., and Mullen, S. L. 2000. Ensemble Forecasting in the Short to Medium Range: Report From a Workshop. *Bull. Amer. Meteor. Soc.* **81**: 2653–2664.

Houtekamer, P. L. and Mitchell, H. L. 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review* **126** (3): 796–811.

Houtekamer, P. L. and Mitchell, H. L. 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review* **129** (1): 123–137.

Jazwinski, A. H. 1970. *Stochastic Processes and Filtering Theory*. Academic Press, New York.

Journel, A. and Huijbregts, C. J. 1978. *Mining Geostatistics*. Academic Press, New York. 600 p.

Kitanidis, P. K. 1995. Quasi-Linear Geostatistical Theory for Inversing. *Water Resour. Res.* **31** (10): 2411–2419.

Li, R., Reynolds, A. C., and Oliver, D. S. 2003. History Matching of Three-Phase Flow Production Data. *SPEJ* **8** (4): 328–340. SPE-87336-PA. DOI: 10.2118/87336-PA.

Liu, N. and Oliver, D. S. 2003. Evaluation of Monte Carlo Methods for Assessing Uncertainty. *SPEJ* **8** (2): 188–195. SPE-84936-PA. DOI: 10.2118/84936-PA.

Liu, N. and Oliver, D. S. 2005a. Critical Evaluation of the Ensemble Kalman Filter on History Matching of Geologic Facies. *SPEREE* **8** (6): 470–477. SPE- 92867-PA. DOI: 10.2118/92867-PA.

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

Lorentzen, R. J., Nævdal, G., Vàlles, B., Berg, A. M., and A-A. Grimstad 2005. Analysis of the Ensemble Kalman Filter for Estimation of Permeability and Porosity in Reservoir Models. Paper SPE 96375 presented at the SPE Annual Technical Conference and Exhibition, Dallas, 9–12 October. DOI: 10.2118/96375-MS.

Nævdal, G., Johnsen, L. M., Aanonsen, S. I., and Vefring, E. H. 2005. Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter. *SPEJ* **10** (1): 66–74. SPE-84372-PA. DOI: 10.2118/84372-PA.

Nævdal, G., Mannseth, T., and Vefring, E. H. 2002. Near-Well Reservoir Monitoring Through Ensemble Kalman Filter. Paper SPE 75235 presented at the SPE/DOE Improved Oil Recovery Symposium, Tulsa, 13–17 April. DOI: 10.2118/75235-MS.

Oliver, D. S., He, N., and Reynolds, A. C. 1996. Conditioning Permeability Fields to Pressure Data. In *European Conference for the Mathematics of Oil Recovery, V* 1–11.

Reichle, R. H., McLaughlin, D. B., and Entekhabi, D. 2002. Hydrologic Data Assimilation With the Ensemble Kalman Filter. *Monthly Weather Review* **130** (1): 103–114.

Ripley, B. D. 1987. *Stochastic Simulation*. John Wiley & Sons, New York.

Skjervheim, J.-A., Evensen, G., Aanonsen, S. I., Ruud, B. O., and Johansen, T. A. 2007. Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter. *SPEJ* 12 (3): 282–292. SPE-95789-PA. DOI: 10.2118/95789-PA.

Tan, T. B. and Kalogerakis, N. 1992. A Three-Dimensional Three-Phase Automatic History Matching Model: Reliability of Parameter Estimates. *Journal of Canadian Petroleum Technology* **31** (3): 34–41.

Tarantola, A. 1987. *Inverse Problem Theory: Methods for Data Fitting and Model Parameter Estimation*. Elsevier, Amsterdam, The Netherlands.

Thulin, K., Li, G., Aanonsen, S. I., and Reynolds, A. C. 2007. Estimation of Initial Fluid Contacts by Assimilation of Production Data With EnKF. Paper to be presented at the SPE Annual Technical Conference and Exhibition, Anaheim, California, 11–14 November.

Tjølsen, C. B., Damsleth, E., and Bu, T. 1994. The Effect of Stochastic Relative Permeabilities in Reservoir Simulation. *Journal of Petroleum Science and Engineering* **10**: 273–290.

Wen, X.-H. and Chen, W. H. 2006. Real-time Reservoir Model Updating Using Ensemble Kalman Filter. *SPEJ* **11** (4): 431–442. SPE-92991-PA. DOI: 10.2118/92991-PA.

Wen, X.-H. and Chen, W. H. 2007. Some Practical Issues on Real-time Reservoir Model Updating Using Ensemble Kalman Filter. *SPEJ* **12** (2): 156–166. SPE-111571-PA. DOI: 10.2118/111571-PA.

Zafari, M., Li, G., and Reynolds, A. C. 2006. Iterative Forms of the Ensemble Kalman Filter. In *Proceedings of the 10th European Conference on the Mathematics of Oil Recovery — Amsterdam* A030.

Zafari, M. and Reynolds, A. C. 2007. Assessing the Uncertainty in Reservoir Description and Performance Predictions With the Ensemble Kalman Filter. *SPEJ* **12** (3): 378–387. SPE-95750-PA. DOI: 10.2118/95750-PA.

Zupanski, M. 2005. Maximum Likelihood Ensemble Filter: Theoretical Aspects. *Monthly Weather Review* **133** (6): 1710–1726.

Dean Oliver is the Mewbourne Chair Professor in Petroleum and Geological Engineering at the U. of Oklahoma, where he has been a faculty member since 2002. From 1997 to 2002, he was on the faculty at the University of Tulsa. He has seventeen years of experience with Chevron prior to joining academia. Oliver is the Executive Editor of SPEJ. He holds a BS degree in physics from Harvey Mudd College and a PhD degree in geophysics from the U. of Washington. **Yaqing Gu** is a reservoir engineer and member of Reservoir Management Team of EPTG, BP America Inc. E-mail: Yaqing.Gu@bp.com. Her working interests include reservoir simulation and characterization as well as uncertainty analysis. She holds a Ph. D. degree in Petroleum Engineering from the University of Oklahoma and B. S. degree in Applied Geophysics from the Guilin Institute of Technology.