The Ensemble Kalman Filter in Reservoir Engineering—a Review

Sigurd I. Aanonsen, Centre for Integrated Petroleum Research, University of Bergen;
Geir Nævdal, International Research Institute of Stavanger; **Dean S. Oliver,** University of Oklahoma;
Albert C. Reynolds, University of Tulsa; and **Brice Vallès,** International Research Institute of Stavanger

Introduction and Background

There has been great progress in *data assimilation* within atmospheric and oceanographic sciences during the last couple of decades. In data assimilation, one aims at merging the information from observations into a numerical model, typically of a geophysical system. A typical example where data assimilation is needed is in weather forecasting. Here, the atmospheric models must take into account the most recent observations of variables such as temperature and atmospheric pressure for better forecasting of the weather in the next time period. A major challenge for these models is that they contain very large numbers of variables.

The progress in data assimilation is because of both increased computational power and the introduction of techniques that are capable of handling large amounts of data and more severe nonlinearities. The aim of this paper is to focus on one of these techniques, the ensemble Kalman filter (EnKF). The EnKF has been introduced to petroleum science recently (Lorentzen et al. 2001a) and, in particular, has attracted attention as a promising method for solving the history matching problem. The literature available on the EnKF is now rather overwhelming. We hope that this review will help researchers (and students) working on adapting the EnKF to petroleum applications to find valuable references and ideas, although the number of papers discussing the EnKF is too large to give a complete review. For practitioners, we have cited critical EnKF papers from weather and oceanography. We have also tried to review most of the papers dealing with the EnKF and updating of reservoir models available to the authors by the beginning of 2008.

The EnKF is based on the simpler Kalman filter (Kalman 1960). We will start by introducing the Kalman filter. The Kalman filter is an efficient recursive filter that estimates the state of a linear dynamical system from a series of noisy measurements. The Kalman filter is based on a model equation, where the current state of the system is associated with an uncertainty (expressed by a covariance matrix) and an observation equation that relates a linear combination of the states to measurements. The measurements are also associated with uncertainty. The model equations are used to compute a forward step (Eqs. 1 and 2) where the state variables are computed forward in time with the current estimate of the state as initial condition. The observation equations are used in the analysis step (Eqs. 3 through 5) where the estimated value of the state and its uncertainty are corrected to take into account the most recent measurements. See, e.g., Cohn (1997), Maybeck (1979), or Stengel (1994) for an introduction to the Kalman filter.

The basic equations for the discrete Kalman filter for a simple linear system are

$$y_n^f = A y_{n-1}^a, \qquad (1)$$

$$C_{y_n^f} = A C_{y_{n-1}^a} A^T + C_\varepsilon, \qquad (2)$$

$$y_n^a = y_n^f + K_n (d_{\mathrm{obs},n} - H y_n^f), \qquad (3)$$

$$K_n = C_{y_n^f} H^T \left(H C_{y_n^f} H^T + C_{d_n} \right)^{-1}, \qquad (4)$$

and

$$C_{y_n^a} = (I - K_n H) C_{y_n^f}. \qquad (5)$$

Here, y represents the estimated state vector for the system; d_{obs} is the observed measurements; C_ε is a covariance matrix representing the model noise; C_y is the covariance matrix of the state vector of the system; C_d is the covariance matrix of the measurement error (which we assume has 0 mean); A is a matrix describing the dynamics of the system; H gives the linear relationship between the measurements and the states; and K is known as the Kalman gain matrix. The subscript n is the time index; the superscript f denotes the outcome from the forward step; and the superscript a denotes the outcome from the analysis step. We have kept the equations simple to avoid introducing unnecessary details at this stage.

The derivation of the Kalman gain matrix (Eq. 4) can be done in several ways. One way of deriving it (see, e.g., Stengel 1994) is by solving a weighted least square problem and imposing the additional constraints that the measurement noise should be independent in time; there should be no correlation between the model noise and the measurement noise. The weighting depends both on the uncertainty in the state variables (C_{y^f}) and the measurements (C_d). Another derivation of the Kalman filter equations can be done based on Bayesian inference (see, e.g., Cohn 1997; Maybeck 1979), which requires the assumptions that both the model noise and measurement noise are Gaussian. Then, all the operations in Eqs. 1 through 5 keep us within the Gaussian probability distributions.

The Kalman filter was extended early to work with nonlinear models through the extended Kalman filter, replacing Eq. 1 with $y_{n+1}^f = F_n(y_n^a)$. Here, F_n is a suitable differentiable function. The extended Kalman filter uses linearizations of the model and observation equations around the estimated mean of the state y_n. The extended Kalman filter gives good performance for a range of systems and was, for instance, used for trajectory estimation in the Apollo space program. The extended Kalman filter has also been used for parameter estimation in hydrological modeling (Eigbe et al. 1998; Eppstein and Dougherty 1996; Hantush and Mariño 1997; Leng and Yeh 2003).

The shortcomings of the extended Kalman filter become severe in cases of highly nonlinear models. Moreover, the equations that need to be solved using the extended Kalman filter become infeasible for large-scale systems because of problems with storage and updating of the covariance matrix C_y.

During the last decades, several alternatives to the extended Kalman filter have been developed. Within the control theory community, several sigma point filters have been introduced (see, e.g., Julier et al. 2000; Lefebvre et al. 2004; Wan and Nelson 2001), primarily for better handling of nonlinear models. The sigma point filters differ from the extended Kalman filter in the way that the uncertainty information is propagated. In the extended Kalman filter, the covariance matrices representing the uncertainty are updated on the basis of the tangent linear models, whereas, in the sigma point filters, the covariance matrices are propagated using sample points to represent the uncertainty and apply forward

Original SPE manuscript received for review 27 March 2008. Revised manuscript received for review 21 December 2008. Paper (SPE 117274) peer approved 1 January 2009.

simulations of each of the sample points with the nonlinear model (or observation) equation. This approach is more robust when the tangent linear model is not a good approximation of the model equations. The number of function evaluations needed for uncertainty propagation using the sigma point filters is proportional to the number of states in the system, and it requires complete storage of the covariance matrix C_y. This means that the sigma point filters break down for large-scale systems.

To handle data assimilation in large-scale systems, the geoscience community has introduced filters where the covariance matrix C_y of the model uncertainty is only partially propagated. The EnKF (see, e.g., Evensen 2003, 2007) and different types of reduced order filters (see, e.g., Rozier et al. 2007) were developed within oceanography and atmospheric sciences during the previous decade. In the forward step, the model equations are applied to each of the ensemble members. The update step is computed using the form of the Kalman filter equation but avoids the computation and storage of the full covariance matrices for the model and measurement uncertainties. The model uncertainty is still formally estimated using the ensemble of state vectors. The initial model realizations used in the EnKF are drawn randomly from the initial distribution, which, to certain degrees, allow us to avoid the limitations of Gaussian assumptions. The analysis step is still based on the first- and second-order moments of the distributions, which, however, limits the performance if the distributions are too far from Gaussian.

Let us mention some other approaches on the basis of reduced order filtering. In these filters, one approximates the covariance matrix C_y by a deterministic selection of eigenvalues and eigenvectors from C_y, as in the SEEK filter (see, e.g., Pham 2001; Rozier et al. 2007), or by combining a deterministic selection with some Monte-Carlo drawing, as in the SEIK filter (see, e.g., Pham 2001). An application of the SEIK filter for updating a reservoir simulation model is presented in Liang et al. (2007). Another approach is based on Karhunen-Loeve decomposition and polynomial expansions to represent random fields. This approach is compared to the EnKF on a 2D synthetic reservoir model (Zhang et al. 2007), but, because the method relies on the ability to represent the random field with a relatively small number of eigenvectors, it is not clear how well it will work on 3D field cases.

The literature on the EnKF is vast, and here we can only point to a few of the most influential papers. The EnKF was first introduced by Evensen (1994). The importance of including random observations to each ensemble member was pointed out by Burgers et al. (1998) and independently included in the formulation used by Houtekamer and Mitchell (1998). The latter paper also argued for the use of covariance localization (limiting the influence of the observations to the state variables that are located spatially close to them) and proposed the use of a pair of ensemble Kalman filters, where one computes the Kalman gain from one ensemble and applies that gain on the other ensemble, and vice versa. In Houtekamer and Mitchell (2001), a practical implementation of covariance localization was suggested using the Schur (elementwise) product. Modifications of the EnKF, avoiding the perturbations of the observations, were suggested by Anderson (2001) and Whitaker and Hammill (2002). Anderson (2001) also included the estimation of a model parameter in an atmospheric model. A review of various issues with the EnKF, including sampling error because of small ensembles, covariance location, filter divergence, and model error, was recently given by Ehrendorfer (2007).

To our knowledge, the first application of the EnKF within the petroleum industry was presented by Lorentzen et al. (2001b). In this paper, model parameters for a dynamic two-phase model of fluid flow in a well were tuned. The method was applied to a set of full-scale experimental data, and improved predictions of the pressure behavior of the well were obtained. This study was followed up by Lorentzen et al. (2003). The number of states in these studies are in the order of a few hundreds, which means that this model can be considered as a medium-size model.

Motivated by the results obtained by Lorentzen et al. (2001b), the EnKF was used to update permeability fields for near-well reservoir models (Nævdal et al. 2002a, 2002b). In Nævdal et al. (2002b), it was shown that improved predictions were obtained when additional data were assimilated; and, in Nævdal et al. (2002a), it was pointed out that there was also an improvement in the quality of the estimated permeability fields as more data were assimilated. After these initial works, the complexity of the studies increased rapidly. In Nævdal et al. (2005), the permeability was estimated on a 2D field-like synthetic example. A series of studies using the PUNQ-S3 model were started with Gu and Oliver (2005). Here, both permeability and porosity were estimated. In Skjervheim et al. (2007), results using the EnKF on real 4D seismic data from a Northern Sea field were presented, the first published demonstration of the use of the EnKF on a field case. The application of the EnKF to field cases will be further discussed in the Field and Pseudofield Examples section.

For non-Gaussian problems, the EnKF does not produce a correct estimate of the posterior probability distribution. This is because the analysis equations are based on information from first- and second-order moments (mean and covariance matrix) only. To get a correct sample of the posterior, other methods are needed. A candidate for getting a better approximation of the posterior distribution is the particle filter (Doucet et al. 2001). In the particle filter, the forward step is computed for a large number of random samples and the posterior distribution after the analysis step is computed by combining the empirical distribution obtained from the samples with the additional information provided from the observations. Applications of particle filters to large-scale systems are nonexistent because of the difficulty of locating particles in regions of high likelihood when the dimension of the model space is large. For small-scale models, some comparisons have been done between the EnKF and particle filters. In Kivman (2003), the bootstrap filter (Gordon et al. 1993) is compared to the EnKF in a parameter estimation problem for a 3D Lorenz system. Using the bootstrap filter gives a better solution in this case, but applying the bootstrap filter to reservoir models still seems out of range with the computer technology of today.

The randomized maximum likelihood (RML) method, introduced independently by Kitandis (1995) and Oliver et al. (1996), can be used to get a correct sample for a linear problem and an approximative solution for nonlinear problems. Zafari and Reynolds (2007) show that, in the limit of an infinite ensemble, the RML and the EnKF are equivalent for linear problems with Gaussian priors. For small, nonlinear test problems with multimodal posteriors, it is also shown that the RML approximately samples the correct posterior, whereas the EnKF fails to capture the multimodal behavior (Zafari and Reynolds 2007). For the PUNQ-S3 reservoir model, the EnKF and the RML were found to give reasonably consistent results (Gao et al. 2006). The RML method needs the computation of gradients, and, with its original implementation, it requires matching of all data at once.

The EnKF is, by its nature, an approach suitable for monitoring. If one is able to monitor the reservoir, a natural follow up is the application of automatic control of the production based on the monitoring. This leads to the field of closed-loop reservoir management. In such cases, one will search for optimal production strategies. It has been shown that this is a promising approach in a number of synthetic waterflooding examples (Brouwer et al. 2004; Nævdal et al. 2006; Overbeek et al. 2004). In these studies, the reservoir model is updated using the EnKF and an optimal flooding strategy is found using adjoint-based optimization. Closed-loop optimization methods that do not require adjoints have been developed by Lorentzen et al. (2006) and Chen et al. (2008). In these approaches, information from the ensemble is used to search for optimal choke settings in waterflooding examples with multisegmented wells. Wang et al. (2007) compare several methods of production optimization, including an EnKF-like method.

Methodology

The EnKF may be introduced in several ways. In the original paper by Evensen (1994), an ensemble of models was introduced as an improvement to the extended Kalman filter by more accurate and efficient estimation of the state covariance matrix. Evensen and van Leeuwen (2000) showed that the general smoother and filter

can be formulated as sequential Bayesian inversion methods, and, in his book, Evensen (2007) gives a thorough presentation of the EnKF as a method for solving the general state and parameter estimation problem. The introduction to the EnKF given here generally follows the description of Evensen (2007) and Skjervheim et al. (2007).

We consider a dynamical model, which is described by a (generally) nonlinear system of partial differential equations, and assume that this equation with boundary conditions has been discretized in space and that the errors in the boundary conditions are 0. Let $u = u(t) \in \mathbb{R}^{N_u}$ denote the discretized approximation to the solution at time t. We further assume that the model depends on some poorly known parameters, $m \in \mathbb{R}^{N_m}$, to be estimated. That is,

$$\frac{du(t)}{dt} = F(u;m) + \varepsilon^m(t), \text{ dynamic model} \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

$$u(t_0) = u_0, \text{ initial state}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

where $\varepsilon^m(t)$ represents the model error. The initial state, u_0 may also be a random variable, representing uncertainty in initial conditions. In a reservoir simulation setting, u typically includes all the primary variables in all grid cells (e.g., pressures and saturations for a black-oil model) at a given time. In parameter estimation applications, it is common to assume that the dominating errors in the model are because of the uncertainty in the parameters, and, in the following, we will neglect the model error, although this is not necessary. Evensen (2007) provides a thorough discussion of model errors in connection with the EnKF and suggests a method for estimating model error as a part of the data assimilation process. Lødøen and Omre (2008) have developed a method for dealing with model bias caused by grid coarseness. In the method of Lødøen and Omre (2008), all model corrections are made relative to a reference, fine scale, while the fluid flow simulations are made on a coarser scale. Each data-assimilation step then consists of three steps. First, a small calibration set of models, both on fine and coarse scale, are used to estimate the dependence between the two scales. Secondly, a large simulation set of coarse-scale realizations is used to predict the forecast on the fine scale based on this dependence. Finally, the measured data are used to update these coarse scale realizations taking into account the bias in the coarse model. Using the estimated dependence between the scales, the large set of model realizations on both fine and coarse scale may then be generated. The methodology is demonstrated on a case study inspired by the characteristics of the North Sea Troll field.

We assume further that the model is constrained by some measurements, collected at times $t_1, t_2, \ldots, t_{N_t}$, and that the measurements are related to the state variables through the (generally) nonlinear relation,

$$d_n = g_n(u(t_n), m) + \varepsilon_n^d \in \mathbb{R}^{N_d}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (8)$$

where $\varepsilon_n^d \sim N(0, C_{d_n})$ represents the measurement error.

The time evolution of the state variable, $u(t)$, which is now a random variable, can be described by the Fokker-Planck or Kolmogorov's equation (Jazwinski 1970). Because we want to condition the solution to the measurements, we consider the model evolution as a sequential Bayesian inversion.

When applying the EnKF to nonlinear problems, it is common to define an augmented state vector containing parameters, state variables, and simulated data, i.e.,

$$y_n = \begin{Bmatrix} m \\ u_n \\ g_n \end{Bmatrix} \in \mathbb{R}^{N_m + N_u + N_d}. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

Note that, with this definition, the relation between the state vector and the data will be given by the linear relation,

$$d_n = H y_n + \varepsilon_n^d, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (10)$$

where H is now a matrix containing only 0s and 1s. Let $y_{n:0}$ denote the sequence $(m, u_0, u_1, \ldots, u_n, g_0, g_1, \ldots, g_n)$, and $d_{n:1}$ denote the sequence $(d_1, d_2, \ldots, d_n)$. The complete solution is then given by the conditional probability density function for $y_{n:0}$ given the data, which can be expressed as,

$$f(y_{n:0} \mid d_{n:1}) \propto f(d_n \mid y_n) f(y_n \mid y_{n-1}) f(y_{n-1:0} \mid d_{n-1:1}), \quad n = 1, 2, \ldots, N_t \quad \ldots\ldots\ldots\ldots\ldots\ldots (11)$$

Details on the derivation of Eq. 11 and the filter equation, Eq. 12, can be found in Evensen and van Leeuwen (2000), Evensen (2007), or Skjervheim et al. (2007). It has been assumed in this derivation that the measurements at a given time are independent of the measurements at other times and that the forward model is a first-order Markov process, which here means that the solution at a measurement time only depends on the solution at the previous measurement time. We say nothing about how the dynamic model is integrated between measurement times, so it is not necessary that observed data exist at every simulator timestep of the numerical solution of Eq. 6. Eq. 11 is a recursive formula, so the solution may be calculated sequentially and forward in time. The first term is the likelihood function for the data at time t_n, the second term represents the forward model, and the last term is the posterior distribution at time t_{n-1}. The product of the last two terms is the prior distribution at time t_n.

The marginal solution at time t_n is the *filter* solution and is obtained by integrating Eq. 11 over $y_{n-1:0}$. The filter solution is given by

$$f(y_n \mid d_{n:1}) \propto f(d_n \mid y_n) f(y_n \mid d_{n-1:1}), \quad n = 1, 2, \ldots, N_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (12)$$

The corresponding solutions for u may be obtained by integrating over m and g, and correspondingly for m. Note that, for $y_{n:0}$ as given by Eq. 11, the solution at all times is conditioned to all data (including measurements at later times). This solution is called the *smoother* solution, as opposed to the filter solution, Eq. 12, which depends only on measurements taken at previous times. The smoother solution may also be obtained by a sequential procedure, such as the ensemble Kalman smoother (EnKS) (Evensen and van Leeuwen 2000; Evensen 2007). Note that the filter and smoother solutions coincide at the time of the last measurement, so the predictions from the smoother and filter solutions will be identical. The EnKS solution may be preferred if the data at different measurement times are correlated, such as difference data in 4D seismic (Skjervheim et al. 2007).

The basic idea of the EnKF is to solve Eq. 12 using Monte Carlo integration in time, where the goal is to sample from the posterior distribution $f(y_n \mid d_{n:1})$. The marginal distributions for u_n and m are then obtained directly from the ensemble for y_n.

Let N_e be the ensemble size. It is convenient to introduce the matrix Y_n holding the ensemble members at time t_n,

$$Y_n = \left\{ y_n^1, y_n^2, \ldots, y_n^{N_e} \right\} \in \mathbb{R}^{N_y \times N_e}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (13)$$

and the matrix $\overline{Y_n}$ storing the ensemble mean in each column,

$$\overline{Y}_n = \left\{ \overline{y_n}, \overline{y_n}, \ldots, \overline{y_n} \right\} = Y_n 1_{N_e}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (14)$$

where 1_{N_e} is the $N_e \times N_e$ matrix with all elements equal to $1/N_e$. Introduce also the ensemble perturbation matrix,

$$\Delta Y_n = Y_n - \overline{Y}_n = Y_n \left(I - 1_{N_e} \right). \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (15)$$

The ensemble covariance matrix can then be expressed as

$$C_{y_n} = \frac{\Delta Y_n \Delta Y_n^T}{N_e - 1}. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (16)$$

We will now state the basic EnKF algorithm.

EnKF Algorithm.

1. Generate N_e independent samples $m_0^{a,j}$ and $u_0^{a,j}$, $j=1,\ldots,N_e$ from the prior distribution of m and the initial distribution of u, respectively.

2. For $n=1,2,\ldots N_t$,

(a) Generate N_e independent, unconditional samples of y_n by integrating each sample from t_{n-1} to t_n using the forward model, Eq. 6. These (unconditional) samples are denoted forecast samples. In y_n, the most updated value for m, denoted m_n, is applied. That is, for $j=1,2,\ldots,N_e$,

$$y_n^{f,j}=\begin{Bmatrix} m_n^{f,j} \\ u_n^{f,j} \\ g_n^{f,j} \end{Bmatrix}=\begin{Bmatrix} m_{n-1}^{a,j} \\ F_n(u_{n-1}^{a,j},m_{n-1}^{a,j}) \\ g_n(u_n^{f,j},m_{n-1}^{a,j}), \end{Bmatrix}, \quad \ldots\ldots\ldots\ldots\ldots\ldots (17)$$

where F_n denotes the forward integration of Eq. 6 from the timestep $n-1$ to n (these timesteps do not have to coincide with the discretization timesteps of the forward model). $y_n^{f,j}$ represents a sample from the prior distribution $f(y_n \mid d_{n-1:1})$. Superscript a denotes the posterior (analyzed) solution from the previous timestep.

(b) Use Monte Carlo integration to estimate the first- and second-order moments from the prior (forecast) using Eqs. 14 and 16.

(c) Generate N_e independent samples from the measurement distribution:

$$d_n^j=d_{\text{obs},n}+\varepsilon_n^{d,j}, \quad j=1,\ldots,N_e. \quad \ldots\ldots\ldots\ldots\ldots\ldots (18)$$

(d) Update each ensemble member using the traditional Kalman filter update, Eqs. 3 and 4. That is,

$$y_n^{a,j}=y_n^{f,j}+C_{y_n^f}H_n^T\left(H_nC_{y_n^f}H_n^T+C_{d_n}\right)^{-1}\left(d_n^j-H_ny_n^{f,j}\right)=y_n^{f,j}+K_n\left(d_n^j-H_ny_n^{f,j}\right), \quad j=1,\ldots,N_e, \quad \ldots\ldots\ldots\ldots (19)$$

where K_n is the Kalman gain matrix, common for all the ensemble members. $y_n^{a,j}$ represents an approximate sample from the filter distribution $f(y_n \mid d_{n:1})$. We see that one complete step of the EnKF consists of a nonlinear prediction and a linear update. It may also be viewed as a sequential simulation with Kriging update and is based on a variance minimizing scheme (Evensen 2007). However, the EnKF is strictly a variance minimizing scheme only when the assumption is made that the prior is Gaussian.

We note that the rank of the covariance matrix, C_y, cannot be larger than N_e-1 and that the full matrix as defined by Eq. 16 never has to be formed when used in the analysis. Thus, the EnKF algorithm is fast, very simple, and does not require much storage, even for a high-dimensional state space.

For a linear model with Gaussian statistics, it can be shown (Evensen 2007) that the analyzed ensemble converges to the correct posterior distribution. Note that a "linear model" here means that both F and g in Eqs. 6 and 8 are linear functions of u and m. Formally, no explicit assumption about Gaussianity or linearity in the forward model is made in the derivation above. The linear relation between the state variable and the data, required for the updating equation, Eq. 19, to be valid, is obtained by the "trick" of adding g_n to the state vector. Li and Reynolds (2007) provide a comprehensive analysis of the EnKF for nonlinear models and consider especially the effect of using the augmented state vector where g_n is included. They show that the EnKF samples correctly from the marginal distributions of m and u provided the forecast distribution for y [i.e., the joint probability density function (pdf) of m, y, and g is Gaussian at each updating step]. This is a strong assumption because it implicitely requires an approximately linear relation between m, u, and g. In practice, the relation between m and g seems to be most critical. This is to be expected because the forward model in a traditional output least squares formulation will be $g(m)$.

Zafari and Reynolds (2007) investigate the EnKF for some nonlinear and multimodal toy problems and show that the basic EnKF fails to capture the structure of the pdf. Still, and despite the simplicity of the method, good results have generally been reported when applying the EnKF to reservoir simulation models. The validity of the EnKF for nonlinear models and non-Gaussian statistics will be thoroughly discussed in the Nonlinear, non-Gaussian section.

The following expressions can be derived for the mean and covariance of the analyzed ensemble, omitting the time index, n:

$$\overline{y^a}=\overline{y^f}+K\left(\overline{d}-H\overline{y^f}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (20)$$

and

$$\begin{aligned}(N_e-1)C_{y^a}&=\Delta Y^a\left(\Delta Y^a\right)^T\\&=\left(Y^f-\overline{Y^f}-KH\left(Y^f-\overline{Y^f}\right)+K\left(D-\overline{D}\right)\right)\\&\left(Y^f-\overline{Y^f}-KH\left(Y^f-\overline{Y^f}\right)+K\left(D-\overline{D}\right)\right)^T\\&=\left((I-KH)\Delta Y^f+KE\right)\left((I-KH)\Delta Y^f+KE\right)^T\\&=(I-KH)\Delta Y^f\left(\Delta Y^f\right)^T-\Delta Y^f\left(\Delta Y^f\right)^TH^TK^T\\&\quad+K\left(H\Delta Y^f\left(\Delta Y^f\right)^TH^T+EE^T\right)K^T\\&\quad+\left((I-KH)\Delta Y^fE^TK^T\right)+\left((I-KH)\Delta Y^fE^TK^T\right)^T. \quad \ldots\ldots (21)\end{aligned}$$

Here, D is the matrix holding the measurement ensemble,

$$D=\left\{d^1,d^2,\ldots,d^{N_e}\right\}\in\mathbb{R}^{N_d\times N_e}. \quad \ldots\ldots\ldots\ldots\ldots\ldots (22)$$

$\overline{D}$ is a matrix with the ensemble mean of d_n^j in each column analogous to Eq. 14, and $E=D-\overline{D}$. We see that, if, in the expression for the Kalman gain (Eq. 4), the ensemble representations are used for C_y^f and C_d, the group of terms on the next to last line of Eq. 21 becomes 0. If, in addition, $N_e\to\infty$, and the distributions of the forecast and observation ensembles are independent, the last line also goes to 0 and the expression for the covariance of the analyzed ensemble reduces to the traditional equation, Eq. 5. For a finite ensemble size, this will not generally be the case, resulting in sampling errors. These sampling errors may be avoided by using a square-root filter.

Using the matrix notation, the analysis equation may be further simplified. Omitting the time index, we obtain (Evensen 2003):

$$\begin{aligned}Y^a&=Y^f+C_{y^f}H^T\left(HC_{y^f}H^T+C_d\right)^{-1}\left(D-HY^f\right)\\&=Y^f+\Delta Y^f(\Delta Y^f)^TH^T\\&\left(H\Delta Y^f(\Delta Y^f)^TH^T+(N_e-1)C_d\right)^{-1}\Delta D\\&=Y^f+Y^f\left(I-1_{N_e}\right)\left(H\Delta Y^f\right)^TC^{-1}\Delta D\\&=Y^f\left(I+\left(H\Delta Y^f\right)^TC^{-1}\Delta D\right)\\&=Y^fX, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (23)\end{aligned}$$

where $\Delta D=D-HY^f$, $C=H\Delta Y^f(H\Delta Y^f)^T+(N_e-1)C_d$, and $X=I+(H\Delta Y^f)^TC^{-1}\Delta D$. We have also used the expression for the covariance matrix, Eq. 16, and that $1_{N_e}\left(H\Delta Y^f\right)^T\equiv 0$. Note that the matrix, X, only depends on the forecast ensemble through the second-order weighted product of $H\Delta Y^f$ and ΔD, so the update can be characterized as a weakly nonlinear combination of the forecast ensemble. If we consider the vector of model parameters, m, it follows from Eqs. 17 and 23 that the solution at all times will be a linear combination of the initial ensemble for m. Thus, it is crucial for the success of the EnKF that the initial ensemble of model parameters properly spans the prior uncertainty.

One useful feature of the EnKF is that forecasts of reservoir performance can be made by running the simulator forward in time from the most recent data assimilation time (analysis step) with each ensemble member. The mean of the ensemble forecast provides an estimate of the true forecast, and the distribution of individual forecasts about the mean provides an estimate of the uncertainty in the estimated forecast. However, forecasts from the last data assimilation may not always be consistent with those obtained by rerunning the simulator from time zero using updated model parameters. Such consistency has only been established for a linear dynamical system with no modeling error where the predicted data is linearly related to the state vector (Thulin et al. 2007). For the nonlinear problems involved in history matching, experiments suggest that this consistency may approximately hold, but it can only be checked by rerunning from time zero using the updated ensemble of model parameters from the final assimilation step. This final rerun from time zero is also advantageous in the sense it eliminates any material balance errors caused by truncating nonphysical values of saturations or porosity to physically feasible values or because of the modification of simulator primary variables with the EnKF analysis step. This rerun from time zero may also be reasonable when the EnKS is used to update initial conditions such as the depths of initial fluid contacts during data assimilation because the updated initial conditions at each data-assimilation step do not influence subsequent predictions forward in time (Evensen et al. 2007; Thulin et al. 2007).

Generation of the Initial Ensemble. Evensen (2004) examines how different sampling strategies and implementations of the analysis scheme influence the quality of the results in the EnKF using a linear advection model. He shows that by first generating a large number of realizations and selecting the initial ensemble from singular vectors associated with the largest eigenvalues of the (approximate) covariance, it is possible to achieve a significant improvement in the EnKF results, using the same number of members in the ensemble. One problem with this is that the property realizations in the initial ensemble will all be relatively smooth and not necessarily consistent with the prior geological model. Evensen (2004) also suggest rescaling the ensemble members to obtain the correct variance. However, it is still not clear how to optimally generate the initial ensemble for a reservoir simulation model. Initial investigations on the effect of initial sampling for reservoir flow problems is presented in Oliver and Chen (2008).

Square-Root Filters. As mentioned above, the basic EnKF algorithm may introduce sampling errors because of the stochastic perturbations added to the observations. However, these sampling errors can be avoided by performing a separate update of the ensemble mean and the ensemble perturbations (Evensen 2004). A number of alternative variants of this, commonly described as ensemble square-root filters (EnSRF), have been proposed, including the ensemble transform Kalman filter (Bishop et al. 2001), the ensemble adjustment filter (Anderson 2001), the local ensemble Kalman filter (Ott et al. 2004) and the ensemble square-root filter (Whitaker and Hamill 2002; Tippett et al. 2003; Evensen 2004). It has been noticed (Lawson and Hansen 2004) that, for simple test models, the ensemble square-root filter may collapse into one state with a single outlier providing the prescribed variance. Leeuwenburgh et al. (2005) discuss the various square-root filters and show that the problem with ensemble collapse may be avoided by a simple randomization of the analysis.

Wang et al. (2004) note that some solutions for generating the ensemble anomalies do not preserve the ensemble mean during the transformation, although they also state that differences are small. Sakov and Oke (2008) carried the analysis further and found that this problem could be significant. As a result, it is recommended that only mean-preserving transformations are used with square-root filters.

In all implementations of the square-root filter, the ensemble mean is first updated using Eq. 20, with $\bar{d}_n$ replaced by the unperturbed measurements. Following Evensen (2007), we then apply the exact version of Eq. 21 (assuming an infinite ensemble size and that the forecast solution and the measurements are independent) together with Eq. 23 to obtain an expression for the ensemble perturbations (again omitting the time index):

$$\begin{aligned}
\Delta Y^a(\Delta Y^a)^T &= (N_e-1)C_{y^a}\\
&= (N_e-1)\left(C_{y^f}-C_{y^f}H^T\left(HC_{y^f}H^T+C_d\right)^{-1}HC_{y^f}\right)\\
&= \Delta Y^f(\Delta Y^f)^T-\Delta Y^f(\Delta Y^f)^TH^T\\
&\left(H\Delta Y^f(\Delta Y^f)^TH^T+(N_e-1)C_d\right)^{-1}H\Delta Y^f(\Delta Y^f)^T\\
&= \Delta Y^f(\Delta Y^f)^T-\Delta Y^f(H\Delta Y^f)^TC^{-1}(H\Delta Y^f)(\Delta Y^f)^T\\
&= \Delta Y^f\left(I-(H\Delta Y^f)^TC^{-1}(H\Delta Y^f)\right)(\Delta Y^f)^T\\
&= \Delta Y^f\left(I-(H\Delta Y^f)^TZ\Lambda^{-1}Z^T(H\Delta Y^f)\right)(\Delta Y^f)^T, \quad \ldots\ldots\ldots(24)
\end{aligned}$$

where $C=Z\Lambda Z^T$ is the eigenvalue decomposition of C. Introducing now the singular value decomposition $\Lambda^{-1/2}Z^TH\Delta Y^f=U_1\Lambda_1V_1^T$, where $U_1\in\mathbb{R}^{N_d\times N_d}$, $\Lambda_1\in\mathbb{R}^{N_d\times N_e}$, and $V_1\in\mathbb{R}^{N_e\times N_e}$, we obtain

$$\begin{aligned}
\Delta Y^a(\Delta Y^a)^T &= \Delta Y^f\left(I-\left(U_1\Lambda_1V_1^T\right)^TU_1\Lambda_1V_1^T\right)(\Delta Y^f)^T\\
&= \Delta Y^f\left(I-V_1\Lambda_1^T\Lambda_1V_1^T\right)(\Delta Y^f)^T\\
&= \Delta Y^fV_1\left(I-\Lambda_1^T\Lambda_1\right)V_1^T(\Delta Y^f)^T\\
&= \left(\Delta Y^fV_1\sqrt{I-\Lambda_1^T\Lambda_1}\right)\left(\Delta Y^fV_1\sqrt{I-\Lambda_1^T\Lambda_1}\right)^T\\
&= \left(\Delta Y^fV_1\sqrt{I-\Lambda_1^T\Lambda_1}\right)V_2^TV_2\left(\Delta Y^fV_1\sqrt{I-\Lambda_1^T\Lambda_1}\right)^T. \quad \ldots\ldots(25)
\end{aligned}$$

In the last line, a random orthogonal matrix product, $I=V_2^TV_2\in\mathbb{R}^{N_e\times N_e}$ is inserted to avoid the ensemble collapse (Leeuwenburgh et al. 2005). It is seen that a possible (randomized) solution for the ensemble perturbations is

$$\Delta Y^a=\Delta Y^fV_1\sqrt{I-\Lambda_1^T\Lambda_1}V_2^T. \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(26)$$

The random orthogonal matrix, V_2^T, may be obtained from a singular value decomposition of a random matrix as stated in Leeuwenburgh et al. (2005) or less expensively from a QR decomposition of an $N_e\times N_e$ matrix of independent random normalized numbers (Evensen 2004). In this decomposition, care has to be taken to get the correct statistical properties (Mezzadri 2007). The eigenvalue decomposition of C may be computationally demanding if the number of measurements is large. However, an efficient subspace pseudoinversion is provided by Evensen (2007).

Sakov and Oke (2008) showed that in order to ensure that the square-root scheme preserves the mean, the symmetric square-root scheme should be used. That is, the product $I=V_1^TV_1$ should be added in Eq. 25, and, thus, in front of V_2^T in Eq. 26. They also gave a procedure for calculating the random matrix, V_2, such that the mean is preserved.

Local Analysis and Large Amounts of Data. If the number of data is large (e.g., when assimilating seismic data) two major problems may occur with the basic EnKF method: (1) It is not computationally feasible to compute or "invert" the matrix $HC_yH^T+C_d$ that appears in the Kalman gain, Eq. 19, or the matrix C in Eq. 23 and (2) as discussed previously, the vector of model parameters, m, in each updated ensemble member must be a linear combination of the corresponding initial ensemble so there are not sufficient degrees of freedom to properly assimilate data when the number of independent data is greater than the number of ensemble members.

The inversion of a large matrix may be avoided by using an ensemble representation of the measurement error covariance matrix, C_d. That is, setting

$$C=H\Delta Y(H\Delta Y)^T+EE^T \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots(27)$$

in Eq. 23 or 24. Inversion of of the $N_d \times N_d$ matrix of Eq. 27 can be replaced by an inversion of a matrix of dimension not larger than N_e using a subspace inversion scheme as noted below. We have seen (Eq. 21), that Eq. 27 also ensures a consistent update of the covariance matrix in the traditional EnKF. However, with this approximation, sampling errors are introduced also in the square-root filter. In addition, a straighforward application of Eq. 27 may result in a loss of rank in the updated covariance matrix (Kepert 2004; Evensen 2007).

One suggestion for dealing with the computationally efficiency and rank issues is to use a subspace EnKF inversion scheme where data mismatches and simulated measurement errors are projected onto the subspace spanned by the principal left singular vectors of $H\Delta Y$ (Evensen 2004, 2007; Skjervheim et al. 2007). There is no assurance, however, that when using this method, one will obtain realizations of model parameters that honor all observed data and, in this case, the estimate of model parameters may be considerably less accurate than the estimate that could be obtained by using a larger ensemble size.

Another procedure that has been tried to attack the problem of assimilating large data sets is local analysis (local updating) (Houtekamer and Mitchell 1998; Skjervheim et al. 2006; Evensen 2007). In the simplest form of local analysis, one would update all the components of the state vector associated with simulator gridblock i using all data in some local neighborhood N_i of gridblock. If the neighborhood includes all gridblocks in the simulation model, then local updating is equivalent to the standard EnKF. If each N_i includes all data that are strongly influenced by the variables of gridblock i, then one would expect local analysis to give reasonable results. However, N_i must be relatively small because we must solve $N\,N_i \times N_i$ matrix problems to perform the local analysis. Here, N is the number of simulation gridblocks. With the standard EnKF update, each analyzed vector $y^{a,j}$, $j = 1,2,\cdots N_e$ is a linear combination of the forecast ensemble. With local updating, if y_i^j denotes the components of y^j that pertain to variables and parameters on gridblock i then each $y_i^{a,j}$ may be updated with a different linear combination of the ensemble of predictions before these components are recombined to obtain $y^{a,j}$. This $y^{a,j}$ is no longer confined to the subspace spanned by the ensemble of predictions, and, thus, with local analysis, it is easier to obtain solutions that match a large data set. Because components are updated independently, the smoothness of an updated state may be less than the smoothness inherent in the covariance of ensemble predictions.

The problems related to the low-rank approximation of the covariance matrix and assimilation of large amounts of data will be further discussed in the following sections.

Covariance Localization

The power of the EnKF is a direct consequence of the ability to approximate the entire covariance matrix from an ensemble that is often orders of magnitude smaller than the number of state variables. As a consequence, however, the straightforward approximation of the covariance suffers from two weaknesses:

1. The rank of the approximation of the covariance is less than or equal to the number of ensemble members, and consequently, the number of perfect data that can be assimilated in a standard EnKF is severely limited. The updates in a standard EnKF are also restricted to the subspace spanned by the members of the forecast ensemble (rank less than or equal to N_e) and the assimilation of a perfect observation removes one degree of freedom from the ensemble (Lorenc 2003). The problem with the rank-deficiency is most significant for problems with large amounts of independent data, such as 4D seismic.

2. Because of the small size of the ensemble from which the covariance is estimated, sampling error can result in spurious correlations that, when used in the Kalman gain, cause changes to the state variables in regions of no real influence. Lorenc (2003) discusses the effect of noise in sampled covariances and shows that, for an ensemble with 100 members, the global variance estimate increased instead of decreased after assimilation of a single perfect observation (one with no error). The conclusion is that the "harm done by spurious covariances with distant grid points is greater than the local benefit."

Although much of the difficulty could be resolved by using larger ensembles, efficiency demands the ability to accurately estimate the forecast covariance from relatively small ensembles.

Motivating Examples and Observations. As seen in Eq. 16, the approximation to the covariance is estimated from the ensemble of realizations of the state vector. Because of the method of approximation, it is clear that the rank of C_y cannot be greater than the rank of ΔY, which is less than or equal to the minimum of $N_e - 1$ and N_y. Also, note that, because the realizations used for the estimate of covariance are random, the estimate is random. **Fig. 1** compares

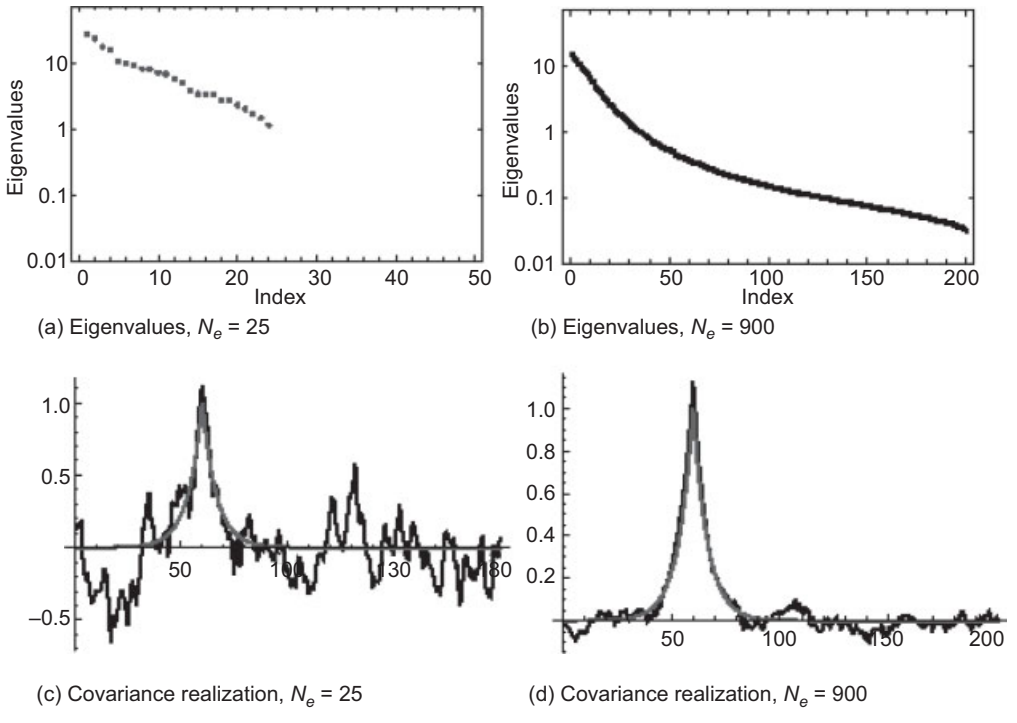


Fig. 1—Comparisons of covariance matrix rank and magnitude of spurious correlations for two different ensemble sizes with N_y = 200.

eigenvalues from covariance estimates generated from an ensemble of size 25 (Fig. 1a) and ensemble size 900 (Fig. 1b). Note that the number of nonzero eigenvalues for the small ensemble is only 24 and that the larger ensemble has 200 (N_y) nonzero eigenvalues. Also, note that the spurious correlations in the covariance estimate from an ensemble size of 25 (Fig. 1c) are much larger than the spurious correlations from an ensemble of 900 (Fig. 1d). The gray curve in both cases indicates the true covariance.

Distance-Dependent Localization. Although it is clear that many of the problems with the rank and erratic nature of the covariance estimate could be eliminated by using a large ensemble, efficiency demands that the size of the ensembles be reduced as much as possible. An effective approach to reduce the influence of spurious correlations on the update is to multiply the ensemble-based estimate of the covariance elementwise by a compactly supported positive/definite matrix ρ to produce a localized covariance estimate $C_{y,\rho} = \rho \circ C_y$.

The elementwise product of matrices is known as a Schur or Hadamard product.

The first application of localization in the EnKF (Houtekamer and Mitchell 1998) simply applied a distance cutoff to the Kalman gain so that only model parameters within a critical distance of the observation were updated. They showed that the optimal radius for the cutoff increased as the size of the ensemble increased.

Houtekamer and Mitchell (2001) describe an efficient implementation of the analysis step that employs a distance-dependent localization of the covariances of the background error calculated from the ensemble using a correlation function having local support to filter the small (and noisy) background-error covariances associated with remote observations. They point out that the method allows the algorithm to function well with relatively small ensembles (even when the number of data is large) and that the updates to the fields are relatively smooth compared to those from methods that simply use a cut-off radius (Houtekamer and Mitchell 1998) or use no localization at all.

Because the covariance matrix for the model and state variables is never formed in EnKF methods, the localization is typically performed on components:

$$[(\rho \circ C_y)H^T][H(\rho \circ C_y)H^T + C_d]^{-1} = [(\rho H^T) \circ (C_y H^T)][(H\rho H^T) \circ (HC_y H^T) + C_d]^{-1}. \quad \ldots\ldots (28)$$

Although the relationship in Eq. 28 is not valid for arbitrary H, it is valid for the block identity matrices typically used in the EnKF for augmented state vectors. Like Houtekamer and Mitchell (2001), most practitioners who use localization appear to have chosen to use the fifth-order compactly supported correlation function of Gaspari and Cohn (1999) for ρ to eliminate spurious correlations.

Hamill et al. (2001) discuss the reasons that distance-dependent filtering of the covariance estimate might be beneficial and investigate the effect of noise in the estimate on the analysis. They also compute the eigenvalues of the estimated covariance matrix with different amounts of localization and show that the covariance matrix becomes full rank when localization is used. This is a consequence of the Schur product theorem (Horn 1990, Page 95, Part c): If matrix B is positive definite and matrix A is positive semidefinite with all of its main diagonal entries positive, then the product $A \circ B$ is positive definite. Because the ensemble covariance estimate is positive semidefinite with positive values on the diagonal, Schur's product theorem guarantees that the Schur product of the ensemble covariance estimate with a positive definite localization matrix will result in a positive definite matrix.

Lorenc (2003) shows that, when the Gaspari-Cohn correlation function is used for localization, the optimal choice for the scale parameter that determines the radius of localization increases as the ensemble size increases. For an ensemble of 100 members, the optimal range is almost twice the (practical) range of the covariance, while, for an ensemble of 1,000 members, the optimal range for localization is approximately three times the covariance range.

Ott et al. (2004) propose local analyses with independent basis vectors determined from singular value decomposition of the local deviations from the mean. From this, they compute a local approximation to the covariance in a subspace spanned by the dominant singular vectors. Because the local basis varies with position, it contains many more degrees of freedom than a global basis, but, like most other localization methods, it ignores the possibility of actual long-range correlations.

Instead of using one particular correlation function for localization of any background covariance, Furrer and Bengtsson (2007) consider the problem of estimating an optimal distance-dependent localization function. This approach involves the minimization of the difference between the true covariance matrix and the localized ensemble estimate of the covariance:

$$\min_{\rho} \mathrm{P} C_Y - \rho \circ C_y \mathrm{P}_E^2, \quad \ldots\ldots (29)$$

where

$$\| C_Y - \rho \circ C_y \|_E^2 = E(\mathrm{tr}((C_Y - \rho \circ C_y)^2)) \quad \ldots\ldots (30)$$

and E represents the expectation. If the true correlation between variables in the state vector is assumed to be a function only of the distance between the variables, then the minimization should be carried out over a class of valid correlation functions that also depend only on distance. They show that, in the isotropic case, the optimal localization functions is

$$\rho(s) = \frac{1}{1 + (1 + p(0)^2 / p(s)^2) / N_e}, \quad \ldots\ldots (31)$$

where, in this equation, $p(s)$ is the forecast covariance function of separation s.

The optimal localization functions determined by this method depend not only on the size of the ensemble but also on the true covariance. **Fig. 2** compares the optimal localization functions for exponential, spherical, and Gaussian covariance functions when the ensemble size is 100.

Combined Ensemble/Spatial Estimation. The update for model parameters is based primarily on the magnitude of the data mismatch and the covariance between predicted data and model parameters. In the EnKF, an estimate of the covariance of the state variable at x' with the state variable at $x'+s$ is obtained from the ensemble of realizations:

$$C_y(x', s) = \frac{1}{N_e - 1} \sum_{k=1}^{N_e} \Delta y_k(x') \Delta y_k(x' + s). \quad \ldots\ldots (32)$$

The error in the estimate of the covariance decreases as $1/\sqrt{N_e}$, so a large ensemble may be necessary for an accurate estimate, and, if the ensemble is small, the error in the estimate may be large (e.g., Fig. 1c).

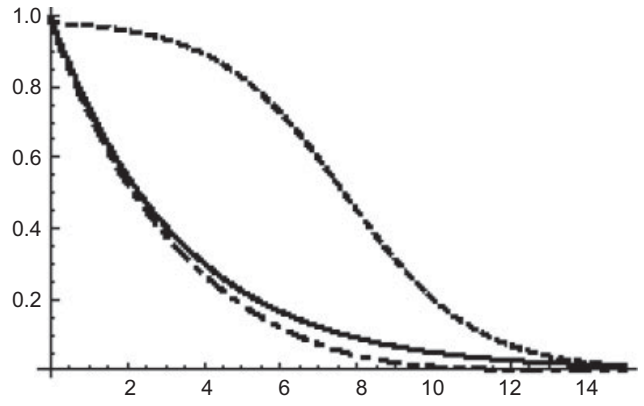
On the other hand, if we assume that the state variables are stationary Gaussian random processes, so that the covariance depends only on the separation of the two variables, then the covariance can be estimated from a single realization (e.g., the kth realization):

$$C_y(s) = \frac{1}{N_s} \sum_{x'} \Delta y^k(x') \Delta y^k(x' + s), \quad \ldots\ldots (33)$$

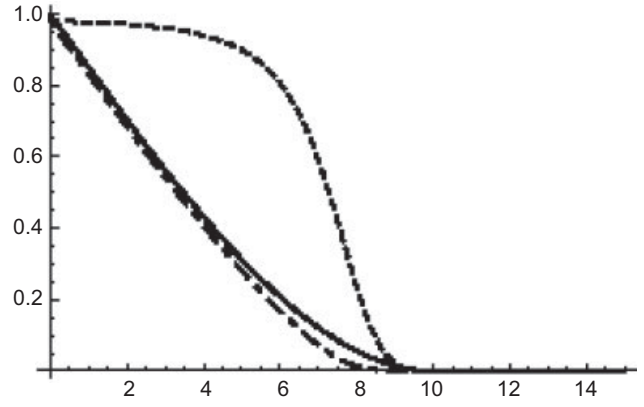
where now the summation is over all N_s pairs that are separated by distance s. Of course, if the process is stationary and there are multiple realizations, it is possible to get smaller errors in the estimate of the covariance by averaging the result from Eq. 33 over all realizations,

$$C_y(s) = \frac{1}{N_s N_e} \sum_{k} \sum_{x'} \Delta y^k(x') \Delta y^k(x' + s). \quad \ldots\ldots (34)$$

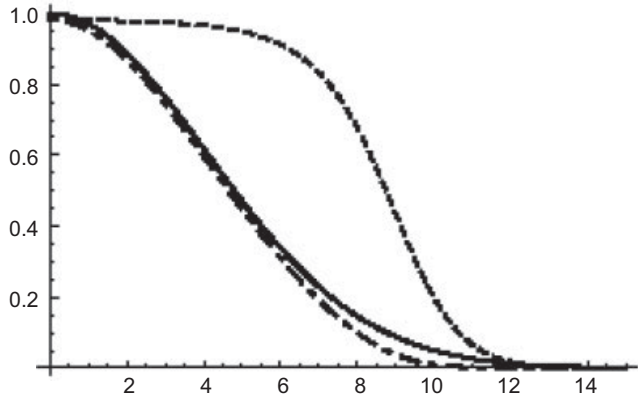
An alternative to purely local (ensemble) or purely global (assuming stationarity) covariance averaging has been described by



(a) Exponential



(b) Spherical



(c) Gaussian

Fig. 2—Localization functions (dashed) estimated using Eq. 31, for three covariance functions (solid). The product of the localization function and the true covariance is also shown (dot-dash). An ensemble size of 100 was assumed.

Buehner and Charron (2007). Their approach can be considered as an average of local spectral and spatial estimates, but it is relatively easy to think of it as an average of shifted realizations such that the correlation estimates are localized.

$$C_{\text{loc}}(x,x') = \frac{\Delta x}{N_e}\sum_k\sum_j \Delta y^k(x+j\Delta x)\Delta y^k(x'+j\Delta x)\mathcal{L}(j\Delta x)\,, \quad \ldots\ldots\ldots (35)$$

where $\mathcal{L}(j\Delta x)$ is a normalized weighting function: $\Delta x\sum_j \mathcal{L}(j\Delta x) = 1$.

The effect of the shift is to increase the effective size of the ensemble.

Pannekoucke et al. (2007) describe the use of a wavelet covariance model for local spatial averaging of wavelet coefficients to improve statistics and achieve a smoother representation of the model parameter correlation. The Schur product was used in conjunction with the wavelet model to reduce spurious long-range correlations.

Outside Information. The methods described in previous sections generally rely on assumptions that the covariance is distance-dependent or that the dependence of the covariance on location is slowly varying so that it can be averaged within a neighborhood. In this section, we cite several other approaches that rely on additional information about the cross-covariance between data and model parameters. The additional information comes either from computation of the sensitivity or influence functions or from knowledge of the true covariance.

Furrer and Bengtsson (2007) consider two types of localization: one that is not distance dependent and one that is model-based and distance dependent. The general idea in both approaches is to reduce the difference between the true forecast covariance matrix and the localized estimate. To do this, they seek a matrix ρ that minimizes the norm of Eq. 29. In the previous section, the localization was assumed to be distance dependent. Furrer and Bengtsson (2007) show, however, that it is possible to minimize Eq. 29 term-by-term, and (ignoring the positive-definiteness constraint) they obtain an expression for each element of the optimal localization:

$$\rho_{ij} = \frac{c_{ij}^2}{c_{ij}^2 + (c_{ij}^2 + c_{ii}c_{jj})/N_e}, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (36)$$

where c_{ij} is the forecast covariance between the ith and jth variables. The potential advantage of Eq. 36 over Eq. 31 is the flexibility it provides for allowing localization functions that do not depend on distance between variables. Examples might include fault transmissibility and original depth of the oil/water contact.

Although the solution they present is not positive definite, it is not restricted to distance dependent localization so it can be applied to nonspatial parameters. One limitation, however, seems to be the need to compute the expectation of the true forecast covariance for observations. For production data, this typically will be expensive.

Like Furrer and Bengtsson (2007), Anderson (2007) proposes a filter for arbitrary localization functions that does not necessarily decrease monotonically with distance from the observation point. The tuning ensemble that is used to estimate localization functions must be several times larger than the normal ensemble. The procedure requires generation of m ensembles of N_e members, where m is a small integer, typically 4. A regression coefficient, β, is estimated for each state variable-observation pair in each ensemble. The mean and the variance of the regression coefficient are used to compute a regression confidence factor (RCF) for each state variable-observation pair. The set of RCFs for each observation form a regression confidence envelope, which can be thought of as a type of localization. One clear advantage of this method is the ability to reduce the effect of spurious correlations on nonspatial parameters such as aquifer strength or initial oil/water contact depth.

Hacker et al. (2007) compare a localization using the Gaspari-Cohn correlation function (Gaspari and Cohn 1999) with optimized range parameter to the hierarchical filter of Anderson (2007). In their test problem (a column model of the atmosphere near the planetary boundary layer), they found that the optimal localization range varied with time of day because of diurnal variation in the amount of mixing. They conclude that localization based on distance alone in a system with multiple data and state types is of questionable value. Bishop and Hodyss (2007) describe a different approach to moderating the influence of spurious correlations on the analysis. In this approach, the moderation functions are generated from powers of the smoothed ensemble correlations. In their experiments, the adaptive moderation gave better results than a nonadaptive moderation when the true error correlation function varied with time and space.

Arroyo-Negrete et al. (2008) and Devegowda (2007) propose a covariance localization method for petroleum reservoir flow that uses sensitivities from the streamline simulation method to quantify the region of influence of model parameters on the observed data. The cross-covariance between production observations and model parameters is assumed to be nonzero only in regions in which the magnitude of the sensitivity (in any of the realizations) exceeds a small threshold value. In application to a synthetic test problem, the method was able to reduce the effect of spurious correlations on the updated permeability fields. Because the localization is based only on the sensitivity function, however, it is not clear how well it can work when the prior covariance has long-range correlations.

Applications of Localization. Localization of the covariance seems to be used in almost all weather prediction applications (Bertino et al. 2003; Anderson et al. 2005; Houtekamer et al. 2005). In the petroleum engineering field, initial applications of localization were primarily to deal with the issue of rank deficiency for problems with large amounts of data. Because localization allows updates to use a different linear combination of realizations for each grid cell, the global update can be obtained from a much larger space than just the span of the ensemble realizations. Both Dong et al. (2006) and Skjervheim et al. (2007) used localization for assimilation of repeated seismic data. Dong et al. (2006) used seismic data at a grid location only to adjust the variables in the data grid column. This ignored much of the spatial continuity constraints in the prior fields. Skjervheim et al. (2007) used a distance-dependent localization that allowed updating over a larger region, but choice of the range was not discussed in the paper.

The other primary reason for localization is to reduce the effect of spurious correlations on the updates of the variables. Devegowda et al. (2007) applied streamline-based analytic sensitivities for covariance localization in the Goldsmith field of Texas. They showed that localization reduced the artificial updating and parameter overshoot at locations far from measurements for an ensemble with 50 realizations.

Agbalaka and Oliver (2008) applied the EnKF to the problem of adjusting facies boundaries in a 3D model. Facies constraints were iteratively enforced using the Gaspari-Cohn distance-dependent localization of the Kalman gain to update the Gaussian random fields for facies. Updates were not applied to the state variables during the iterations.

Non-Gaussian Priors and Strongly Nonlinear Data Functions

The EnKF works quite well when the prior probability distributions are Gaussian and when the relationships between the model parameters, state variables, and observation variables are approximately linear. In a comparison of several different algorithms for data assimilation reported by Verlaan and Heemink (2001), it was concluded that the EnKF is to be preferred for strongly nonlinear problems. Still, violation of one or both of these two assumptions may introduce severe problems when using the EnKF for data assimilation in multiphase flow petroleum reservoirs. In this section, we will examine separately approaches for dealing with nonlinearity in the dynamical system and the complex non-Gaussian distributions for model parameters (e.g., reservoir models characterized by a facies distribution).

State Variables. Two-phase flow in porous media generally exhibits behavior in which the saturation of the displacing phase is high at locations behind the saturation "front" and low ahead of the front. When the location of the front is uncertain, the pdf for saturation in the region of the front is bimodal (Chen et al. 2009). Use of the standard EnKF analysis formula (Eq. 4) for saturations can result in saturations that exceed physically possible limits (Gu and Oliver 2006). In practice, this does not always seem to be a problem, as a number of field cases have been successfully carried out using standard EnKF methods with simple truncation of implausible values (see the Field and Pseudofield Examples section). The importance of the effect of nonlinear/non-Gaussian behavior has been observed to be important, however, when forecasts from the ensemble are much different than the observations (Gu and Oliver 2007) and when the geology is complex (Agbalaka and Oliver 2008). In these cases, the analysis from the traditional EnKF methods can be extremely poor.

Methods for dealing with nonlinearity in the state variables for ensemble methods can be categorized into two main approaches: parameterization of the state variables and iterative filters. We address each of these approaches in turn.

Parameterization. Because the EnKF analysis step performs poorly when the prior pdf for variables are not Gaussian, it is natural to suppose that applying a transform to the non-Gaussian variables might improve the performance of the filter.

Bertino et al. (2003) transformed concentration of phytoplankton (nonnegative) in an ecological oceanographic model to a more nearly Gaussian variable by means of a log-normal transformation. They note that, although there is no easy check to determine if the variables that were transformed were nearly multivariate Gaussian, the filter performance was much improved by the transformation. Gu and Oliver (2006) applied a similar approach to water saturation in a 1D waterflood using a normal score transform. The transformed variables used in the EnKF helped avoiding unphysical values for water saturations sometimes observed with untransformed variables. However, unrealistic oscillations between high and low water saturation values were not always prevented.

Chen et al. (2009) reparameterize the state vector, replacing a non-Gaussian variable (water saturation) with another related state variable (saturation arrival time) that is approximately Gaussian. Although the pdf for saturation is observed to be bimodal in some regions, the location of the saturation front along a streamline and the time of arrival of a saturation at a particular location are both approximately Gaussian. Note that, in streamline simulation, the time-of-flight variable is used to parameterize location along a streamline (Datta-Gupta and King 1995) and, in traditional history matching, the arrival time of saturation is sometimes used as an observation (Wu and Datta-Gupta 2002). The primary differences of Chen et al. (2009) from the traditional EnKF are the need to estimate arrival time and the requirement to interpolate from arrival time curves to saturation.

This method (Chen et al. 2009) is developed for fluid displacement problems and follows on the work published by Gu and Oliver (2006) who replaced water saturations in the state vectors with a geometrical variable locating the position of the shock front. The results for a 1D test case improved the water saturation profiles, but an extension to two or three dimensions did not appear to be practical. Chen et al. (2009) proposed to use the water front arrival time instead of front location because of the ease of adaptation to a finite difference simulator and because arrival time is thought to be quasilinearly related to the petrophysical properties of the reservoir (Wu and Datta-Gupta 2002). The results obtained by Chen et al. (2009) show improvements over the traditional EnKF for the field water saturation distribution.

Jafarpour and McLaughlin (2007a, 2007b) introduce the use of the discrete cosine transform (DCT) parameterization method to the problem of history matching of reservoir models. DCT finds its origins in signal and image processing and is the compression method used in JPEG image compression and MPEG video compression standards (Rao and Yip 1990). DCT is a Fourier related transform that uses orthonormal cosine basis functions for representing an image that, in our case, can correspond to the spatial distribution of state variables or model parameters. The powerful compression property of DCT allows for retaining only a few (r) basis functions relatively to the total number of gridblocks (n). Jafarpour and McLaughlin (2007a, 2007b) showed that a 5% r/n ratio was sufficient to give a reasonably good approximation. Note that DCT is one example of decomposing a data set into basis functions and that one could theoretically use other principal component analysis (PCA) methods that have more physically oriented base functions but could require more computational time.

Jafarpour and McLaughlin (2007b) combine the EnKF with the DCT. Eq. 19 is modified such that the coefficients of the retained

cosine basis functions representing m and u_n are updated instead of these quantities. The updated model parameters and state variables are then constructed from these updated coefficients.

This approach has been tested on two 2D, two-phase, synthetic reservoir models, where DCT parameterization was applied to both state variables (pressure and saturation) and model parameters (permeability). Note that one of the test cases was a two-facies reservoir model thereby indicating that DCT can help solving problems encountered in complex geology. The results obtained by Jafarpour and McLaughlin (2007b) are as good or slightly better than the traditional EnKF method. The DCT parameterization, by retaining only a few basis functions, emphasizes large-scale events (channels for instance) thereby improving connectivity and limiting spurious correlations with the advantage that it reduces the computational cost of the EnKF update. This can be of importance when working on very large reservoir models in real-time application.

Iterative Ensemble Filters. One advantage of the EnKF compared to other methods is its computational efficiency. State variables and model parameters can often be updated simultaneously without the need to rerun the model.

Occasionally, however, the problem is so nonlinear that the EnKF computes updated state variables and model parameters that are obviously unphysical. In certain cases, this problem is overcome by truncation to avoid unphysical values. In other cases, more sophisticated alternatives are required as iterative methods.

Most iterative forms of the EnKF can be viewed as algorithms for minimizing a stochastic objective function or, equivalently, for maximizing the posterior probability of each realization conditional to its forecast value and the new data. Several of these methods (Reynolds et al. 2006; Gu and Oliver 2007; Li and Reynolds 2007) can be viewed as ensemble approximations of the RML method, which is known to sample correctly for linear problems with Gaussian priors. Another iterative form of the EnKF can be viewed as an ensemble approximation to the iterated extended Kalman filter (Lorentzen and Nævdal 2008) where only the EnKF analysis step is repeated. In addition, Wen and Chen (2006, 2007) have proposed methods that primarily seek to preserve the physical plausibility of the state variables through iteration.

Iterative Ensemble Maximum Likelihood Filter. Zupanski (2005) proposes an ensemble maximum likelihood filter for general nonlinear observations. The ensemble in this method is used to approximate the Hessian and the gradient of the objective function in the subspace spanned by the ensemble, and the iteration is performed only on the analysis step. The author suggests that it can be viewed as a maximum likelihood approach to the ensemble transform filter (Bishop et al. 2001).

Iterative Ensemble Filters Based on RML. Because some methods seek only to estimate the model parameters (and perhaps initial conditions) and use the reservoir simulator to solve the dynamical equations to forecast the state variables while other iterative forms attempt to estimate both the model parameters and state variables at some earlier time then forecast to the present, the objective function that is minimized varies somewhat among the methods.

Reynolds et al. (2006) and Li and Reynolds (2007) describe an iterative method that attempts to generate realizations of the reservoir model parameters m at timestep t_n by solving the iteration equation

$$m_{k+1}^{a,j} = m_k^{a,j} - \left[C_{m^f} - C_{m^f,d_n^f}(C_{d_n} + C_{d_n^f,d_n^f})^{-1} C_{d_n^f,m^f} \right] \times \left\{ C_{m^f}^{-1}(m_k^{a,j} - m^{f,j}) + G_{m^f,j,k}^T C_{d_n}^{-1}\left(g_n(m_k^{a,j}) - d_n^j\right) \right\}, \quad \ldots\ldots (37)$$

where k is the iterative index. Here, C_{m^f} is the prior covariance matrix for $m^f = \left\{ m^{f,1}, \ldots, m^{f,N_e} \right\}$ estimated over all ensemble members as in the EnKF method. The matrix $C_{d_n^f,d_n^f}$ is the covariance matrix between the simulated measurements at t_n (denoted by d_n^f), and C_{m^f,d_n^f} is the covariance matrix between m^f and d_n^f. Note that $C_{m^f,d_n^f} = C_{d_n^f,m^f}^T$. The term $G_{m^f,j,k}$ defines the derivative of $g_n(y_{n,k}^j)$ evaluated at $y_{n,k}^j$, the jth column vector of $Y_{n,k}$ as defined in Eq. 13. The second term in curly brackets in Eq. 37 is a gradient calculated for each ensemble member by using adjoint methods as in the RML method (Reynolds et al. 2006), while all covariance matrices are calculated in a similar manner as in the EnKF method. In each iteration, it is necessary to perform one forward simulation from time zero to the current assimilation timestep and one solution of the adjoint system backward from the current assimilation timestep to time zero to compute the gradient of the objective function at each iteration timestep (Li and Reynolds 2007).

Li and Reynolds (2007) also propose a method that does not require the simulator to be rerun from the initial conditions. In this method, iteration is used to improve the estimate of the model and state variables at the previous assimilation time, conditional to observations at the current assimilation time. Like the previously discussed iterative methods that rerun from the initial time, this method is known to produce a valid sampling of the posterior pdf for linear problems with Gaussian priors. For nonlinear approximations, it is relatively inexpensive because the adjoint solution for gradient computation is only necessary on the time interval back from the current assimilation time to the previous assimilation time. Data assimilation results were not as good, however, as results from the method that iterated over the entire time interval.

The methods of Reynolds et al. (2006) and Li and Reynolds (2007) use adjoint methods to compute the gradients required for minimization. In general, a separate gradient is computed for each realization and each realization can use a different iteration step size for each iteration. During iteration, the ensemble is primarily used to compute an approximation to the Hessian. Because each ensemble member uses a separate gradient computed locally, the method is able to sample multiple peaks in a pdf from a single ensemble of realizations. It should also be expected that minimization in the neighborhood of a minimum for adjoint-based methods should be more rapid than for methods that use an ensemble-average gradient. The iterative EnKF based on RML proposed by Gu and Oliver (2007) differs from the two previous works in that it uses, in the iterative formula, the ensemble to compute an approximation to the average sensitivity matrix and the preconditioning for minimization. As a result, it does not require the use of an adjoint code to compute gradients, but it then uses one descent direction for all realizations. It would not be expected to sample multimodal distributions well, but it might be better than adjoint-based methods for problems with multiple local minima (Annan and Hargreaves 2004; Liu and Oliver 2005a).

Because the dimension of the parameter or state space is so large in most real data-assimilation problems, it is difficult to know the shape of the objective function in real problems. Projections of the objective function onto a low dimensional space for two single-phase flow problems are shown in **Fig. 3.** The objective function of Fig. 3a, from a reservoir in which permeability is a Gaussian random field, is differentiable and has two minima of approximately equal depth. If this were the appropriate objective function, it would be important for a sampling method to draw samples from both minima. Results from the EnKF, or an iterative form of the EnKF that used a single Kalman gain for all realizations, would do a poor job of representing the true pdf for this problem, while RML or an iterative EnKF method that used different gradients for each realization (Li and Reynolds 2007) would be expected to sample the pdf well. Perhaps for a similar reason (that an average gradient may give poor results for problems with multiple minima) Zang and Malanotte-Rizzoli (2003) reported that, for a multimodal distribution, the error of the ensemble mean for the EnKF using larger ensembles can be larger than with smaller ensembles. The objective function on the right (Fig. 3b), which resulted from a facies model, is not differentiable everywhere and has multiple local minima (Zhang et al. 2003). An iterative EnKF method that uses an average gradient computed from the ensemble (Gu and Oliver 2007), may be better at avoiding local minima in this type of problem.

Iterative Ensemble Filter Based on the Iterated Extended Kalman Filter. Lorentzen and Nævdal (2008) propose an iterative EnKF approach motivated by the iterated extended Kalman filter (Lefebvre et al. 2004; Jazwinski 1970) in which the measure-

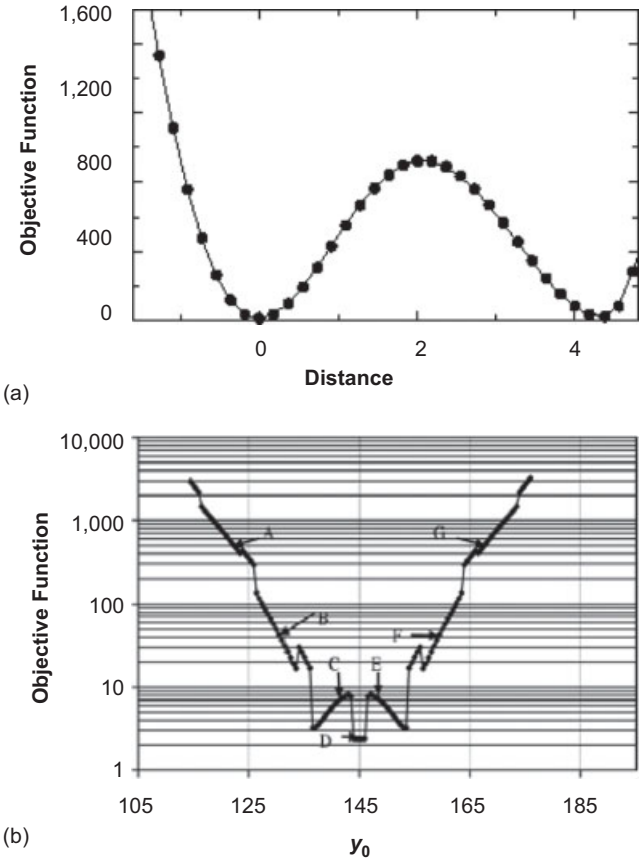


Fig. 3—Objective functions for single-phase flow in a Gaussian random field with pressure data at a producing well (a) and for single-phase flow in a reservoir with adjustments to location of facies boundary (b). Fig. 3b is from Zhang et al. (2003).

ment model is linearized around the analyzed state estimate. This approach relies on using a given *observational model*, g, expressing the simulated measurements as function of the updated model parameters and state variables to produce updated and improved simulated measurements. The iteration procedure applies to the analyze step of the EnKF method only. In their method, only one Kalman gain matrix is estimated for the ensemble at each iteration step.

This approach has been tested on simple mathematical models and the results obtained outperform the traditional EnKF method for nonlinear problems. As with the traditional EnKF, this method seems easy to implement with any reservoir flow simulator software. However, the downside of this approach depends on the observational model that is usually not known for reservoir models.

Iterative Ensemble Filter for Plausibility. Other approaches (Wen and Chen 2006, 2007; Gu and Oliver 2006) consist of iteratively updating the estimate of model parameters then using the simulator to forecast from the last assimilation time t_{n-1} with the old state variables (obtained at t_{n-1}) and new model parameters (obtained at t_n). The method ensures physically plausible state variables, but, despite some improvements in the values of state variables and model parameters on 1D and 2D synthetic reservoir models, this approach is difficult to justify because it fails to reduce to the standard EnKF for linear problems (Gu and Oliver 2007; Zafari and Reynolds 2007). In particular, the method updates model parameters but not the state variables at the previous assimilation time.

Model Parameters for Complex Geology. Most current reservoir models built by geoscientists are characterized by a distribution of facies, which can be defined as a distinctive body of rock with specified characteristics such as mineralogy, sedimentary structures, grain size, or fossil content (Cross and Homewood 1997). Consequently, a reservoir facies model is a structural description of a reservoir in terms of 3D objects, each corresponding to a facies type. Like petrophysical properties, facies can often be inferred from well logs at well locations, but the spatial distribution of facies between wells is highly uncertain. To distinguish the different facies, it is common to assign them numerical labels. These numbers are labels only, however, not numerical values and cannot be represented by Gaussian distributions.

For the EnKF to become a useful tool in updating and history matching realistic reservoir models, it is necessary to show that it can be applied to reservoir models whose petrophysical properties are not adequately modeled as Gaussian random fields. The approach currently taken is to attempt to identify a suitable manner to "transform" the different facies types into intermediate random variables having Gaussian distributions so that the EnKF methodology can be applied successfully. It would appear also that generating realistic initial ensemble members is of great importance when dealing with reservoir facies models as suggested by Jafarpour and McLaughlin (2009). These authors demonstrated that remark by generating their initial ensemble members using multipoint statistics (Strebelle 2002).

Four general approaches to addressing the issue of the EnKF for complex geology can be distinguished thus far: the truncated pluri-Gaussian (Liu and Oliver 2005a, 2005b; Agbalaka and Oliver 2008; Zhao et al. 2008), the level set method (Moreno and Aanonsen 2007), the Gaussian mixture models (Dovera and Della Rossa 2007), and the discrete cosine transform (Jafarpour and McLaughlin 2007b) discussed earlier in the Parameterization section.

Truncated Pluri-Gaussian. Liu and Oliver (2005a, 2005b) introduced the use of truncated pluri-Gaussian method (Galli et al. 1993) as a tool for updating reservoir facies models by means of the EnKF. A concise chronological review of the method development can be found in Liu and Oliver (2005b). The papers by Agbalaka and Oliver (2008) and Zhao et al. (2008) extend the application to a 3D reservoir and greater nonlinearity in the production data.

The method as usually applied is based on the truncation of two Gaussian random fields (GRFs) Z_1 and Z_2 that are defined on the entire reservoir domain. The facies at any location in the reservoir can be determined from the values of Z_1 and Z_2 and a truncation map that assigns regions of the Z_1-Z_2 domain to a particular facies. Liu and Oliver (2005a) suggested using the GRFs Z_1 and Z_2 in the state vector instead of the petrophysical properties because the assumptions of a Gaussian prior are more likely to be satisfied with this approach. The vector of model parameters, m, in Eq. 9 then consists of Z_1 and Z_2 in each grid cell. To honor well observations, the measurement vector, d_n^j, also includes the simulated facies observations at the well locations. The issue of handling nonnumerical facies data in the updating procedure is solved by replacing the simulated facies types in d_n^j by the facies mismatch, $f.f$ is equal to 0 when the simulated and real observed facies types are identical and is equal to 1 otherwise.

Agbalaka and Oliver (2008) show that using localization through a Schur product acting on the Kalman gain matrix reduced the need for very large ensemble size and the risk of excessive loss of variability in the ensemble, as previously observed (Liu and Oliver 2005a, 2005b).

Agbalaka and Oliver (2008) propose to decouple the assimilation of production and facies data where production data are assimilated first and facies data afterward. An iteration loop on facies data only, inspired by (Liu and Oliver 2005a, 2005b), ensures that perfect match for the facies types at well locations is achieved. The decoupling formulation is used to avoid improper weighting of the production data that could lead to large changes in the updated state variables and/or destroy previously matched facies data. This decoupling implies that the vectors d_n^j contain only the facies mismatch part during the iteration on the facies constraints. Zhao et al. (2008) use a different approach to ensure that ensemble members honor facies observations at wells. Essentially, their procedure consists of redoing the EnKF analysis step using pseudodata for the facies data whenever the normal proposed EnKF analysis step does not honor hard data. These pseudodata represent data for the pair of GRF values derived to attempt to find

a realization of the facies distribution that honors facies data but is as close as possible to the corresponding realization that would be obtained without using pseudodata in the EnKF update equation.

Regardless of the method used to incorporate facies data, the results are very satisfactory as the updated reservoir facies representations and the facies proportions are in agreement with the reference case (see **Fig. 4** for examples of reservoir models after assimilation of production data). Moreover, the authors showed that running the reservoir flow simulator from time zero with the latest updated model parameters gives better forecast results than when performing the forecasting directly from the last assimilation timestep. It could be because the state variables are reinitialized, thereby avoiding possible inconsistency with the updated model parameters (Agbalaka and Oliver 2008). On the other hand, for a truncated pluri-Gaussian model that yields facies distributions similar to those shown in Fig. 4, Zhao et al. (2008) and Zhao (2008) found that predictions of production data forward from the final data assimilation time were quite consistent with predictions obtained by rerunning the simulation from time zero using the ensemble obtained at the final data-assimilation step. They also found for the same example that reasonable realizations of the permeability and porosity for each facies could be obtained coincident with updating the GRFs with the EnKF, however, they assumed that each facies is homogeneous and that permeability is isotropic.

Disadvantages of the truncated pluri-Gaussian method include difficulty in determining truncation maps and structural properties of the GRFs that are suitable for highly complex reservoirs such as river mouth, meanders, or carbonate systems. Secondly, although the underlying model parameters are multivariate Gaussian, the relationship between observations and model parameters is highly nonlinear and iterative methods may frequently be required for data assimilation.

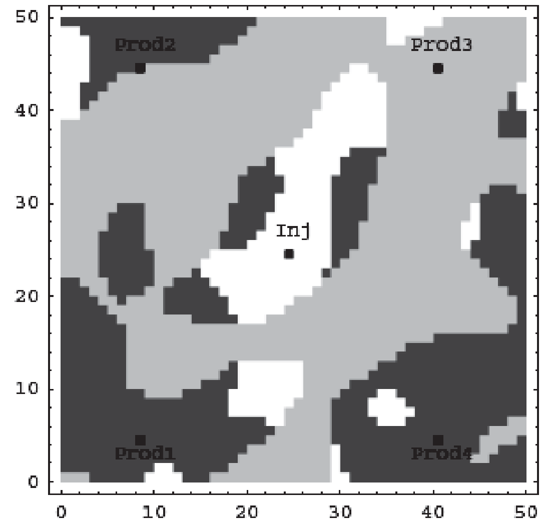
Level Set Method. The level set method was first introduced by Osher and Sethian (1988) for studying surface motion problems such as crystal growth or flame propagation. Since its inception, this method has been applied to many fields, including image processing, computational geometry, and computational fluid dynamics. One application of importance to petroleum geoscientists and engineers is the mapping of 3D objects from seismic data (Richardsen and Randen 2005). This mapping of objects, or seismic facies, coupled with geological information, obtained from well-logs for instance, can be used to generate reservoir facies models. Recently, variants of the level set method have been applied to history matching (Lien et al. 2005; Berre et al. 2007; Villegas et al. 2006a, 2006b; Nielsen 2006).

For any closed curve $\Gamma(\tau)$ in 2D, the level set function $\phi(x(\tau),\tau)$ is usually defined as the signed distance from the curve $\Gamma(\tau)$; it is positive for any points enclosed by $\Gamma(\tau)$, negative for any points outside $\Gamma(\tau)$, and equal to 0 for all points on $\Gamma(\tau)$ (Sethian 1999). The symbol τ corresponds to the time variable used in the discretization of the equations of the level set method. The motion of $\Gamma(\tau)$ is matched with the zero level set of the level set function, $\phi(x(\tau),\tau)=0$, and the resulting initial value partial differential equation for the evolution of the level set function resembles a Hamilton-Jacobi equation (Sethian 1999):

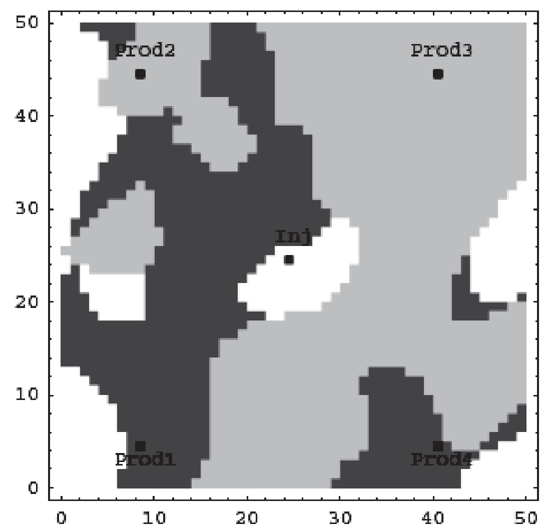
$$\frac{\partial\phi}{\partial\tau}+V|\nabla\phi|=0. \quad \text{.................................(38)}$$

The "velocity" V indicates how, in what direction, and at what speed, the interface $\Gamma(\tau)$ should be relocated. It will guide the movement of this interface from its initial position toward the true boundary. One of the crucial parts of the level set method is then a proper choice for V that would reveal the important information for edges and connectivity (Richardsen and Randen 2005).

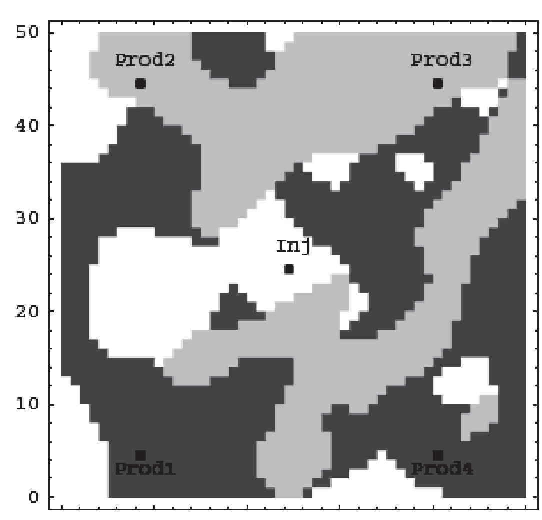
In the context of reservoir modeling, Γ might represent the boundary between two different facies types. If more than two facies are present, then it may be necessary to define multiple level set functions—n level set functions can be used to represent up to 2^n subdomains. A piecewise constant permeability field can then be expressed as a sum of basis functions whose boundaries can be estimated by minimizing data mismatch.



(a) Reference field



(b) Realization 4



(c) Realization 15

Fig. 4—Facies maps from Layer 2 for ensemble sizes 240 with localization of facies adjustment (Agbalaka and Oliver 2008).

Moreno and Aanonsen (2007) propose combining the level set method with EnKF. Here, Eq. 38 is used to evolve the level set function $\phi(x(\tau),\tau)$ such that it minimizes data mismatch. The

velocity V at which the level set evolves is modeled as a GRF and is the variable to be estimated with the EnKF. Different realizations of velocity give different reservoir facies models. The authors have tested their approach on a 2D synthetic reservoir model composed of two facies types. Their ensemble of 74 realizations was updated sequentially in time using the EnKF where bottomhole pressures and water cuts were the assimilated data. The final updated realizations presented developed strong similarities with their reference case, and the final facies proportions of the updated fields are in close agreement with the reference as well. It is, however, not yet clear how easily constraints from complex reservoir models can be captured by the method.

Gaussian Mixture Models. Because the distribution of petrophysical properties in reservoirs with distinct facies are typically multimodal, Dovera and Della Rossa (2007) have proposed a new form of the EnKF that can be applied to sample from a Gaussian mixture model (GMM). Previous works have used GMM but in the context of particle filtering for oceanic or atmospheric data assimilation problems (Anderson and Anderson 1999; Hoteit and Pham 2008). Gaussian mixture models are models in which the pdfs of model parameters are parametrically described as weighted sums of Gaussian pdfs (Hastie et al. 2001). Dovera and Della Rossa (2007) show that, for linear inverse problems with GMM as a prior, the posterior density is also GMM. To make the EnKF formulation suitable for GMM, Dovera and Della Rossa (2007) propose a new analysis step scheme in which they derive expressions for the conditional means, conditional covariances, and weights for the mixture model. The authors also provide a synthetic example of nonlinear data assimilation for a one-parameter model with a bimodal prior. The bimodal posterior density is correctly sampled by the GMM form of the EnKF, while the traditional EnKF results in a single mode.

Field and Pseudofield Examples

Overview. Despite outstanding theoretical and practical questions on the applicability of the EnKF for highly nonlinear and non-Gaussian problems—how to include modeling error, how best to control spurious correlations caused by sampling error, how best to control filter divergence caused by a limited ensemble size, and how to estimate the measurement errors and eliminate outliers—applications of the EnKF to both pseudofield cases (Nævdal et al. 2005; Gu and Oliver 2005; Lorentzen et al. 2005; Gao et al. 2006) and real field cases (Skjervheim et al. 2007; Haugen et al. 2008; Evensen et al. 2007; Bianco et al. 2007) have been very encouraging. Perhaps the most important observation is that, in each of these field cases, a significantly better match of production data was obtained with the EnKF than was obtained using manual history matching. Skjervheim et al. (2007) considered assimilating production data together with seismic data, whereas Haugen et al. (2008), Evensen et al. (2007), and Bianco et al. (2007) considered only production data.

Although it is well understood that model error should be included in the forecast step, a rigorous procedure to properly include modeling error has not been established for multiphase flow in porous media, and, hence, in the field examples discussed here, modeling error was either ignored or given such a small variance about its mean of 0 that it had no effect on the results.

Perhaps motivated by the results of Houtekamer and Mitchell (1998), who found an ensemble size of 128 sufficient to obtain accurate results for a large scale atmospheric physics problem, examples in the reservoir engineering literature, including the field examples discussed in this section, have frequently used an ensemble size on the order of 100. There is no guarantee that this is always adequate, and the number necessary to avoid problems (spurious correlations, insufficient degrees of freedom to assimilate a large number of independent data, underestimation of uncertainty) could be larger in some cases. Whatever the ensemble size, one should recognize that, at the end of each data-assimilation step (EnKF analysis step), each realization of the vector of model parameters is a linear combination of the initial realizations as shown in the Methodology section, unless some type of localization is used in the analysis step. Thus, to obtain a good approximation of the true vector of model parameters, the truth must be well approximated by a vector in the subspace spanned by these initial ensemble members. Note that the EnKF shares this feature with the gradual deformation method (Hu et al. 1999) although the EnKF typically uses a much larger number of realizations.

Wen and Chen (2006, 2007) considered the effects of ensemble size for a 2D waterflooding problem on a $50 \times 50 \times 1$ uniform grid with a water injection well near the center of the reservoir and a producing well near each of the four corners. The permeability field of the true reservoir is such that the water injection well and one production well both lie in a connected zone of high permeability, whereas low-permeability barriers fall between the injector and two of the other producers. They considered ensemble sizes of 50, 100, 200, and 400. Their results (Wen and Chen 2006, Fig. 8) indicate that the mean log-permeability field computed from the 50-member ensemble case gives a poor estimate of the truth, but the 100-member ensemble gives a reasonable estimate of the truth. The ensemble mean log-permeability computed for ensemble sizes of 100, 200, and 400 are not radically different. However, the map of gridblock variances of the log-permeability field indicates that the variances tend to increase significantly with the ensemble size. Even the variance for the 400-member ensemble is noticeably larger than the corresponding variance for the 200-member case at many of the simulator gridblocks. They showed (Wen and Chen 2006, Fig. 9) that predictions of bottomhole pressure were biased when predictions were done with ensemble sizes of 50 and 100. For the problem they considered, one is forced to conclude that an ensemble size on the order of 200 is required to accurately characterize the uncertainty in the estimate of the mean and in reservoir performance predictions. The results of Wen and Chen (2006) are consistent with the analysis of van Leeuwen (1999), who showed that the EnKF systematically underestimates the variance for small ensembles.

One way to test whether 100 realizations is sufficient to characterize uncertainty using the EnKF is to compare statistical estimates obtained from a suite of ensembles with each ensemble consisting of 100 realizations generated from the prior pdf. Lorentzen et al. (2005) have done this for the well known PUNQ-S3 example (Floris et al. 2001), a small synthetic reservoir that encapsulates the geology and operating conditions of a real field that contains six producing wells. Lorentzen et al. (2005) considered 10 ensembles, each of which contained 100 independent realizations generated from the prior model specified by Floris et al. (2001) and conditioned to well observations of permeability and porosity. (They actually considered two different prior models, but that is not germane to this discussion.) They applied the Kolmogorov-Smirnov test to reject the hypotheses that the 10 CDF's reflect the same distribution of forecasts. Although the authors did not investigate this aspect further, they did note that "it is tempting to suspect that it is related to the limited ensemble size."

Because the number of seismic data will be far larger than the number of ensemble members, matching seismic data poses special problems and, hence, will be discussed independently.

Assimilating Production Data. ***Field Example 1.*** Haugen et al. (2008) used the EnKF to improve the data mismatch obtained by manually history matching production data from a North Sea field. The reservoir model is based on a $45 \times 75 \times 26$ grid with 45,000 active cells. The model parameters are the cell values of horizontal log-permeability and porosity. Each cell vertical permeability is updated by setting it equal to 0.1 times the updated value of cell horizonal permeability. No other model parameters were included, and the initial pressure and saturation distributions, including location of the initial fluid contacts, were assumed to be known. The reservoir structure was also assumed to be known. They used an ensemble consisting of 100 realizations with the initial ensemble of model parameters generated using a Gaussian variogram in each layer where each layer variogram was derived from the geological model. The geological model was on the same grid as the reservoir simulation model so upscaling was not required. The production data consisted of bottomhole pressure (WBHP), gas/oil ratio (WGOR), oil production rate (WOPR), and water cut (WWCT), with wells produced at the historical specified

total rate during the data-assimilation step with a bottomhole pressure constraint equal to atmospheric. On average, production data measured at the four producers were assimilated twice a month during a five-year period. They truncated any negative saturations obtained during data assimilation to 0 and truncated saturations greater than 1 to 1.

The probability distribution for measurement errors is needed to compute the data covariance matrix involved in the Kalman gain and is also needed to generate the perturbations of observed data. Haugen et al. (2008) specified the standard deviations for measurement errors and assumed that measurement errors were uncorrelated. They also implemented an ad hoc data filtering scheme to reduce the effect of outliers. When a predicted datum deviated so much from the corresponding observed datum that it was classified as an outlier, the standard deviation of the measurement error for that datum was increased so the datum is weighted less in the EnKF analysis scheme.

Their results indicate that the specified standard deviations of measurement error affect the results (Fig. 9 of their paper), but, even with significant variation in the specified measurement error, the EnKF still gives a better match of data during the data-assimilation process than was obtained by manual history matching. With the base case specification of measurement errors, the match of production data during assimilation with the EnKF is significantly better than was obtained manually. Rerunning the simulation from time zero with the updated permeability and porosity fields, which appear to be reasonable, gives a better match than was obtained by manual history matching but noticeably worse than was obtained during the data-assimilation step of the EnKF. They note without demonstration that, when they did not use the measurement filter, they encountered instabilities in some of the simulations. The effect of not using any filter had far more effect on the results than varying the specified standard deviations of measurement errors.

Field Example 2. Bianco et al. (2007) applied the EnKF to assimilate production data from the Zagor field, which is an onshore field in west Africa. The east part of the field has a primary gas cap on top of the oil rim, whereas the west part of the field contains only gas underlain by water. The initial conditions, including the positions of the initial gas/oil and gas/water contacts (established from geological/geophysical data) were assumed known. The field consists of a 32-foot-thick sand body with seven layers of good quality sands with porosity between 0.1 and 0.24. The reservoir simulation grid is $156 \times 77 \times 10$ but contains only 25,669 active cells. The field contains eight appraisal wells plus two producing wells; all wells are fully penetrating.

On the basis of core and log data at the 10 wells, the authors set vertical permeability equal to horizontal permeability and constructed a deterministic relation to calculate horizontal permeability from porosity. Thus, the only reservoir parameters that were included in the state vector during data assimilation were the porosities of the active gridblocks. Each layer porosity field was represented by a 2D isotropic variogram with a very long correlation range (4000 m). The variance ranged from 0.07 in the top layer to 0.1 in the bottom layer. With this information, initial realizations of the gridblock porosities used in the initial ensemble of state vectors was generated by sequential Gaussian simulation conditioned to the hard data for porosity at the 10 wells. The EnKF was then used to assimilate WWCT and WGOR data at 33 timesteps during a three-year period with each timestep approximately equal to 30 days. Only one of the two wells produced water. No model error was added to predictions (forecasts) from the reservoir simulator. Analyzed saturations less than 0 were set equal to 0, and those greater than 1 were truncated to 1. As in the field case presented by Haugen et al. (2008), the authors found data quality to be an issue. For data considered to be unreliable, the standard deviation of the measurement error was set equal to 600% of the measured value. Thus, the effect of unreliable data on the EnKF analysis step is negligible because the measurement error for reliable water-cut data were assigned a standard deviation of 0.05, and the standard deviation of the WGOR measurement error ranged from 0.2 to 0.6 Mscf/STB. The maximum WGOR observed was approximately 12 Mscf/D. During data assimilation, the measured monthly oil rates were set as the target rates. It appears that, during the simulation runs, these target rates were satisfied.

Bianco et al. (2007) generated results for three different ensemble sizes, 50, 100, and 135. They found that simulation runs using the initial ensemble with no assimilation of production data gave poor matches of the production data. A much improved data match was obtained using the EnKF. They quantified the goodness of match using the average data mismatch function defined by

$$J_A = \frac{1}{2N_e} \sum_{j=1}^{N_e} \sum_{i=1}^{N_d} \left(\frac{d_{\text{obs},i} - d_i^j}{\sigma_i} \right)^2, \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (39)$$

where N_d is the total number of data, $d_{\text{obs},i}$ is the ith measured data, d_i^j is the corresponding predicted data obtained by rerunning from time zero with the jth updated ensemble member of rock properties, and σ_i is the standard deviation of the measurement error associated with $d_{\text{obs},i}$. Although this is not exactly the customary root mean square of the data mismatch terms, one might still expect to reduce the value of J_A below $N_d = 99$ if the standard deviations are correctly specified and we obtain a correct sampling of the posterior pdf conditional to all data. Using values of d_i^j generated from the forward run of the initial ensemble without conditioning to production data gave values of J_A ranging from 1,136 to 1,197, depending on the number of ensemble members (50, 100, or 135), whereas the value of J_A calculated by running the simulator from time zero using the final set of ensemble members generated with the EnKF gave values of J_A ranging from 273 to 308. Thus, conditioning to the production data with the EnKF decreased the value of J_A by almost a factor of 4. In the 50- and 100-ensemble member cases, the estimate of porosity fields (the ensemble mean) from the last data-assimilation step contained some abnormally high porosity values (greater than 0.42) with most, but not all, of this overshooting occurring in the west part of the reservoir where no well data were available. In the 135-ensemble member case, all ensemble average gridblock porosities were below 0.34.

Field Example 3. Evensen et al. (2007) applied the EnKF to assimilate production data from a North Sea reservoir. The black-oil simulation model has approximately 82,000 active gridblocks. The reservoir has four production wells and one water injection well. The basic structural model of the field and the locations of faults and geological layers are assumed to be known sufficiently accurately so they may be fixed during data assimilation. Model parameters include fault transmissibilities, gridblock horizontal log-permeability and porosity, vertical transmissibility multipliers, and the depths of initial fluid contacts. Vertical permeability is fixed equal to 10% of horizontal permeability. The state vector consists of 328,000 dynamic variables plus the reservoir parameters, which are gridblock porosities and horizontal log-permeabilities, 42 fault transmissibilities, 24 vertical transmissibilities, and the depths of five initial water/oil and gas/oil contacts. The depth of each fluid contact was modeled as a Gaussian random variable with mean equal to the best-guess value and standard deviation equal to 20 m. The initial ensemble was conditioned on log-data. An ensemble with 100 members was used. The wells were produced by specifying the total rate of withdrawal at reservoir conditions.

The EnKF was used to assimilate oil rate, gas rate, and water-cut data at each well. The solid curve in **Fig. 5** represents the estimated depth (mean depth computed from the realizations) of the water/oil contact in one region of the reservoir at each data-assimilation step, and the error bars represent ±3 standard deviations. Note that the estimated depth of the water/oil contact and its uncertainty change very little up to approximately 600 days, suggesting the depth is uncorrelated to production data. The estimated depth (continuous curve) decreases from approximately 600 to approximately 1,200 days with little change in its uncertainty. Then, shortly after 1,200 days, the estimate of WOC depth increases very quickly and there is a significant reduction in uncertainty. As production well P3 begins production at 1,200 days, it is possible that the large change in the estimated depth shortly after 1,200 days is related to bringing well P3 online. Although the depths of the initial water/oil contacts were updated at each data-assimilation step, these updated values

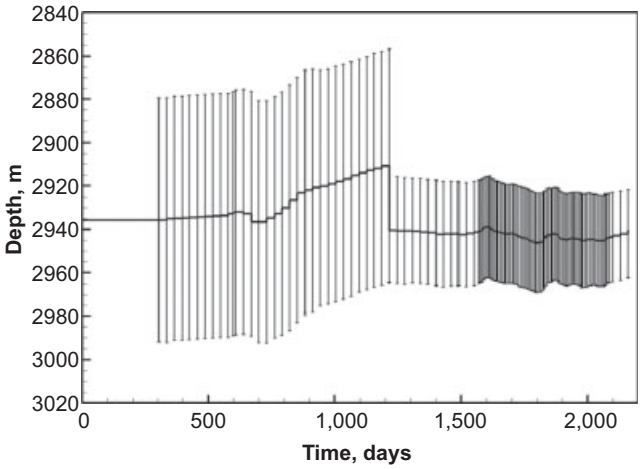


Fig. 5—Evolution of estimated water/oil contact (courtesy of Geir Evensen).

are not used at subsequent reservoir simulation timesteps and thus do not explicitly affect these future predictions. On the other hand, the initial values of fluid contact depths determine the initial saturation distributions and the initial and analyzed saturation values strongly affect subsequent predictions.

Evensen et al. (2007) reported that good matches of data were obtained during data assimilation but showed only the production data generated by running the reservoir simulator from time zero using the final ensemble obtained by assimilating all production data with the EnKF. The authors show clearly that the EnKF provides an improved reservoir model that predicts data that match the measurements much better than the predictions from the initial ensemble. The results also show that a reduction in uncertainty is achieved by integrating the production data using the EnKF. However, as shown in **Fig. 6,** the EnKF is unable to assimilate the late-time producing GOR data at well P1 correctly and predictions from the EnKF ensemble give significantly early breakthrough times at well P2, although the later-time water-cut data at well P2 are matched fairly well. In Fig. 6, the thinner black curves represent predictions from the first 20 members of the initial ensemble and the wider black curves represent results obtained by running the simulator from time zero using the first 20 ensemble members obtained by assimilating all observed data (solid black circles). Whether the inability to match some data is because of excessive noise in the data, model error, the exclusion of important parameters, or something else is unclear. For example, relative permeability curves were assumed known but relative permeabilities can significantly influence the water cut and producing gas/oil ratio.

Assimilation of Seismic Data. In assimilating 3D or 4D seismic data, or interpreted seismic data such as acoustic impedance or Poisson's ratio, we effectively have a data measurement corresponding to each gridblock. Thus, in assimilating seismic data, the matrix that must be inverted to apply the Kalman gain to the data mismatch term is extremely large. Moreover, to introduce a sufficient number of degrees of freedom to match seismic data accurately may require a significantly larger ensemble size.

Skjervheim et al. (2007) used a subspace EnKF inversion scheme to integrate interpreted seismic data into simulation models for both a 2D synthetic model and a real field case. When 4D seismic data are assimilated, one must use the ensemble Kalman smoother (EnKS) because 4D data represents the difference between seismic attributes at the time of the first and second seismic surveys. Thus, at every data-assimilation step in the period between the surveys, one must update the state vector at the first survey time to obtain the correct forecast of the difference data. Between the times of the two seismic surveys, we assume we have production data to assimilate.

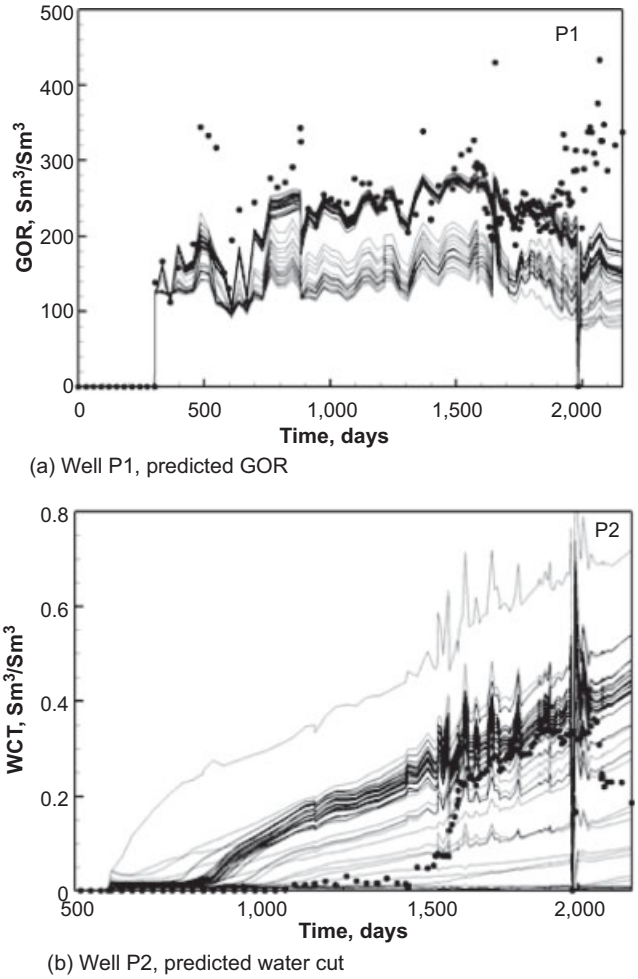
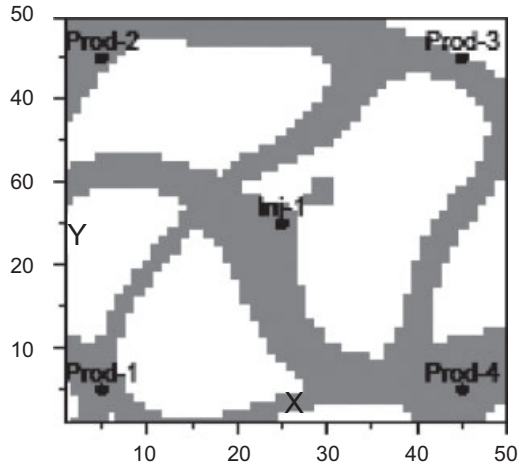
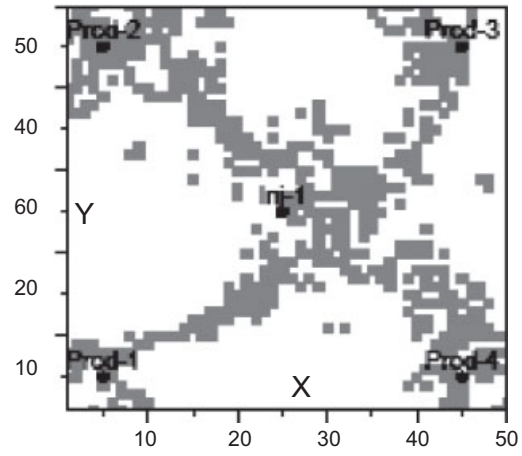


Fig. 6—Production predictions from rerun and initial ensemble (first 20 realizations shown) (courtesy of Geir Evensen).

In the 2D synthetic example, the simulation model is defined on a $15 \times 15 \times 1$ grid with a water injection well in one corner and a producing well in the opposite corner. The truth case was generated with a different variogram than was used to generate the initial ensemble of the permeability field, resulting in initial realizations that were much smoother than the truth. An ensemble size of 100 was used. The EnKF was used to estimate gridblock permeabilities by assimilating bottomhole pressures, producing GOR, and water-cut data with all other reservoir properties fixed at the true values. In most cases, they also assimilated inverted seismic data. Production data were available every three months for seven years, and, after simulating data, they made forward predictions up to 10 years. They also made predictions from time zero with the estimated permeability field (the ensemble mean) to compare with predictions with the true model. When only production data were assimilated, good data matches were obtained but the ensemble mean permeability did not capture the trends even though forward predictions were reasonable and unbiased. In the two cases where production data in combination with the difference between acoustic impedance data sets collected at five and seven years were used, the correlation coefficient between the estimated and true permeability fields were on the order of 0.4, the lowest of all the cases where seismic data were used as conditioning data. Moreover, in these two cases, which differ only in the magnitude of the standard deviation of the measurement error that was specified, the rerun from time zero with the average field did not match the flowing wellbore pressure data even though predictions forward from the last data-assimilation step were reasonable and unbiased. Interestingly, when the two seismic data sets at five and seven years were assimilated individually, instead of simply using



(a) True facies distribution



(b) Realization from global/local analysis

Fig. 7—Comparison of true facies distribution with a realization obtained by EnKF with a global/local analysis (courtesy of Yong Zhao).

difference data, the correlation coefficient between the estimated and true permeability fields increased to 0.6, predictions from the end of data assimilation were unbiased with lower uncertainty, and the run from time zero with the ensemble mean permeability is almost identical to predictions from the truth.

Although the authors give no explanation of why difference data gives worse results than assimilating seismic data directly, Zhao et al. (2008) suggest two possibilities. First, the variance of the difference data is twice the variance of the individual data sets. Second, using difference data requires updating of the state vector at the time of the first seismic survey several times as production data measured between the times of the two surveys are assimilated. This could allow errors caused by strong nonlinearities or non-Gaussianity to accumulate. On the other hand, it is not clear that the authors included uncertainty in static parameters in the rock physics model, and the main reason for using difference data is to attempt to cancel the effect of inaccurately known model parameters that influence seismic data.

Skjervheim et al. (2007) also assimilate a field seismic data set with the EnKF. In this field example, they assimilate Poisson's ratio difference data together with production data to improve a base case model for the reservoir. The reservoir model consists of four partially communicating fault blocks and contains two water injection wells, one gas injector, and three producing wells. The base case permeability field was significantly modified by assimilating production data, and the data match was improved; but, the reduction in the root mean square error compared to the base case was only 14% when only production data were assimilated and was approximately 25% when both production and seismic difference data were assimilated. As the authors note, on the basis of their 2D synthetic case, it is possible that better matches would be obtained by assimilating the two Poisson's ratio data sets individually instead of using difference data. The authors also note that the code they used assumes seismic data measurement errors are spatially uncorrelated, which is unlikely to be true. In fact, Skjervheim et al. (2006) show a 3D synthetic example where the performance of the EnKF is severely degraded when one uses a diagonal covariance for seismic data measurement error when the true measurement error is spatially correlated over five gridblocks horizontally and two gridblocks in the vertical direction.

Skjervheim et al. (2006) present 2D and 3D synthetic cases where they assimilate both production and time-lapse seismic data to estimate permeability, porosity, and clay ratio. They compare global analysis (both seismic data and production data are updated globally using the standard EnKF analysis equation) and global/local analysis where production data are assimilated globally using the standard EnKF analysis equation but time-lapse seismic data are assimilated locally. To assimilate seismic data globally, they use a subspace inversion scheme. The 2D example they consider is based on a 15×15 grid. Despite this small size, the global scheme yields poor results when an ensemble size of 100 is used; forward predictions of production data are biased, and the correlation coefficient between the true permeability field and the estimated permeability field (the ensemble mean) is only 0.37. Increasing the ensemble size to 250 and using global analysis gave a correlation coefficient between mean and estimated permeability fields of 0.64 as compared to coefficients of 0.78 using global/local analysis with an ensemble size of 100, and 0.81 for global analysis with 800 ensemble members. However, the data match, as evaluated by the root mean square error, was larger (182) for the global/local method than for the global method with an ensemble size of 250 (160) or an ensemble size of 800 (155). These last three methods all gave reasonable unbiased predictions of production data, although the spread of predictions with the 250-ensemble case is noticeably smaller than the spread of predictions from the 800-ensemble member case and the global/local 100-ensemble member case. The permeability field obtained with the global/local analysis procedure was somewhat less smooth than obtained from all the global analysis cases.

For some problems, local analysis can result in a field that loses much of the smoothest inherent in the prior geological model. **Fig. 7** shows results obtained by Zhao et al. (2008). In this work, the authors assimilated production and seismic data with the EnKF to update a pair of Gaussian random fields used in the definition of a truncated Gaussian model (Galli et al. 1993; Liu and Oliver 2005b) for a geological system consisting of two facies. The plot on the left is the truth, which was used to generate the synthetic data, and the plot on the right is a realization obtained by assimilating data with a global/local analysis EnKF scheme. Note that the true model roughly represents a channel system, but, as shown by the right plot, the continuity of the channel is lost because the local updating does not enforce the continuity inherent in the prior model.

Conclusions and Outlook

We have presented a comprehensive review of the application of the EnKF within petroleum engineering and updating of reservoir simulation models in particular. To our knowledge, the first paper appeared in 2001. Since then, the popularity of the method has been steadily growing, and the total number of publications by the end of 2007 exceeded 40. Despite the limitations of the methodology, many of the results presented in the literature have been very encouraging.

The main challenges when using the EnKF to update reservoir simulation models are related to the low rank representation of the model covariance matrix, non-Gaussian prior models, strong nonlinearities in the forward model, and the application to large-scale field models. These issues have been considered in detail. Another important issue, which has not been much considered yet, is model error.

Effective updating of reservoir models and prediction of future performance with uncertainty continues to be a major challenge for the industry, and the idea of continuously updating the models as new data becomes available without having to rerun from time zero is very attractive, especially in a closed-loop reservoir management setting. Although many challenges remain, the EnKF appears to be one of the most promising methods presently available.

Nomenclature

A = linear model operator
C = covariance matrix
D = matrix containing measurement ensemble
d = vector of measurements (data)
$E = D - \overline{D}$ = matrix of measurement perturbations
$E[\cdot]$ = expected value
F = nonlinear model operator
$f(\cdot)$ = probability density function
G = sensitivity matrix
$g(\cdot)$ = forward model (predicted data)
H = measurement operator
I = identity matrix
$J(\cdot)$ = objective function (data mismatch function)
K = Kalman gain matrix
m = vector of model parameters
N_d = number of measurements (data)
N_e = ensemble size (number of ensemble members)
N_m = number of model parameters
N_t = number of assimilation times
N_u = dimension of state variable
$p(s)$ = covariance function of separation s
t = time
u = dependent (state) variable
X = ensemble updating matrix
x = spatial variable
Y = matrix containing ensemble of state vectors
$\overline{Y} = Y1_{N_e}$ = matrix containing the ensemble mean in each column
y = EnKF state vector
Z = Gaussian random field
Γ = facies boundary
$\Delta D = D - HY$ = ensemble of innovation vectors
$\Delta Y = Y - \overline{Y}$ = ensemble perturbation matrix
ε = stochastic error vector
ρ = covariance localization function
τ = level set parameter
ϕ = level set function
$1_{N_e} = N_e \times N_e$ = matrix with all elements equal to $1/N_e$

Subscripts

e = ensemble
k = iteration index
n = time index
obs = observed

Superscripts

a = analyzed (a posteriori)
d = data
f = forecast (a priori)
j = ensemble member index
m = model
T = matrix transpose

Acknowledgments

The authors acknowledge financial support from the Norwegian Research Council, PETROMAKS program. The third author acknowledges the support provided by the OU Consortium for Ensemble Methods. The work of the fourth author was also partially supported by the member companies of The University of Tulsa Petroleum Reservoir Exploitation Projects and the US Department of Energy under Award No. DE-FC26-04NT15517. However, any opinions, findings, conclusions or recommendations herein are those of the authors and do not necessarily reflect the views of the Department of Energy. The authors would also like to acknowledge the reviewers and the editor for their valuable comments.

References

Agbalaka, C. and Oliver, D.S. 2008. Application of the EnKF and Localization to Automatic History Matching of Facies Distribution and Production Data. *Mathematical Geosciences* **40** (4): 353–374.

Anderson, J.L. 2001. An Ensemble Adjustment Kalman Filter for Data Assimilation. *Monthly Weather Review* **129** (12): 2884–2903. doi: 10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

Anderson, J.L. 2007. Exploring the Need for Localization in Ensemble Data Assimilation Using a Hierarchical Ensemble Filter. *Physica D: Nonlinear Phenomena* **230** (1–2): 99–111. doi: 10.1016/j.physd.2006.02.011.

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. doi: 10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2.

Anderson, J.L., Wyman, B., Zhang, S., and Hoar, T. 2005. Assimilation of Surface Pressure Observations Using an Ensemble Filter in an Idealized Global Atmospheric Prediction System. *Journal of the Atmospheric Sciences* **62** (8): 2925–2938. doi: 10.1175/JAS3510.1.

Annan, J.D. and Hargreaves, J.C. 2004. Efficient Parameter Estimation for a Highly Chaotic System. *Tellus A* **56** (5): 520–526. doi: 10.1111/j.1600-0870.2004.00073.x.

Arroyo-Negrete, E., Devegowda, D., Datta-Gupta, A., and Choe, J. 2008. Streamline-Assisted Ensemble Kalman Filter for Rapid and Continuous Reservoir Model Updating. *SPE Res Eval & Eng* **11** (6): 1046–1060. SPE-104255-PA. doi: 10.2118/104255-PA.

Berre, I., Lien, M. and Mannseth, T. 2007. A Level-Set Corrector to an Adaptive Multiscale Permeability Prediction. *Computational Geosciences* **11** (1): 27–42. doi: 10.1007/s10596-006-9037-3.

Bertino, L., Evensen, G., and Wackernagel, H. 2003. Sequential Data Assimilation Techniques in Oceanography. *International Statistical Review* **71** (2): 223–241.

Bianco, A., Cominelli, A., Dovera, L., Nævdal, G., and Vallès, B. 2007. History Matching and Production Forecast Uncertainty by Means of the Ensemble Kalman Filter: A Real Field Application. Paper SPE 107161 presented at the EAGE/EUROPEC Conference and Exhibition, London, 11–14 June. doi: 10.2118/107161-MS.

Bishop, C.H. and Hodyss, D. 2007. Flow-Adaptive Moderation of Spurious Ensemble Correlations and its Use in Ensemble-Based Data Assimilation. *Quarterly Journal of the Royal Meteorological Society* **133** (629): 2029–2044. doi: 10.1002/qj.169.

Bishop, C.H., Etherton, B.J., and Majumdar, S.J. 2001. Adaptive Sampling With the Ensemble Transform Kalman Filter. Part I: Theoretical Aspects. *Monthly Weather Review* **129** (3): 420–436. doi: 10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Brouwer, D.R., Nævdal, G., Jansen, J.-D., Vefring, E.H., and van Kruijsdijk, C.P.J.W. 2004. Improved Reservoir Management Through Optimal Control and Continuous Model Updating. Paper SPE 90149 presented at the SPE Annual Technical Conference and Exhibition, Houston, 26–29 September. doi: 10.2118/90149-MS.

Buehner, M. and Charron, M. 2007. Spectral and Spatial Localization of Background-Error Correlations for Data Assimilation. *Quarterly Journal of the Royal Meteorological Society* **133** (624): 615–630. doi: 10.1002/qj.50.

Burgers, G., van Leeuwen, P.J., and Evensen, G. 1998. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review* **126** (6): 1719–1724. doi: 10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Chen, Y., Oliver, D.S., and Zhang, D. 2008. Efficient Ensemble-Based Closed-Loop Production Optimization. Paper SPE 112873 presented at the SPE/DOE Symposium on Improved Oil Recovery, Tulsa, 21–23 April. doi: 10.2118/112873-MS.

Chen, Y., Oliver, D.S., and Zhang, D. 2009. Data Assimilation for Nonlinear Problems by Ensemble Kalman Filter With Reparameterization. *Journal of Petroleum Science and Engineering* **66** (1–2): 1–14. doi: 10.1016/j.petrol.2008.12.002.

Cohn, S.E. 1997. An Introduction to Estimation Theory. *Journal of the Meteorological Society of Japan* **75** (1B): 257–288.

Cross, T.A. and Homewood, P.W. 1997. Amanz Gressly's Role in Founding Modern Stratigraphy. *Geological Society of America Bulletin* **109** (12): 1617–1630. doi: 10.1130/0016-7606(1997)109<1617:AGSRIF>2.3.CO;2.

Datta-Gupta, A. and King, M.J. 1995. A Semianalytic Approach to Tracer Flow Modeling in Heterogeneous Permeable Media. *Advances in Water Resources* **18** (1): 9–24. doi: 10.1016/0309-1708(94)00021-V.

Devegowda, D., Arroyo, E., Datta-Gupta, A., and Douma, S.G. 2007. Efficient and Robust Reservoir Model Updating Using Ensemble Kalman Filter With Sensitivity Based Covariance Localization. Paper SPE 106144 presented at the SPE Reservoir Simulation Symposium, Houston, 26–28 February. doi: 10.2118/106144-MS.

Dong, Y., Gu, Y., and Oliver, D.S. 2006. Sequential Assimilation of 4D Seismic Data for Reservoir Description Using Ensemble Kalman Filter. *Journal of Petroleum Science and Engineering* **53** (1–2): 83–99. doi: 10.1016/j.petrol.2006.03.028.

Doucet, A., de Freitas, N., and Gordon, N. eds. 2001. *Sequential Monte Carlo Methods in Practice*. New York: Statistics for Engineering and Information Science, Springer.

Dovera, L. and Della Rossa, E. 2007. Ensemble Kalman Filter for Gaussian Mixture Models. In *Petroleum Geostatistics 2007, 10–14 September 2007, Cascais, Portugal*, A16. Utrecht, The Netherlands: Extended Abstracts Book, EAGE Publications BV.

Ehrendorfer, M. 2007. A Review of Issues in Ensemble-Based Kalman Filtering. *Meterologische Zeitschrift* **16** (6): 795–818. doi: 10.1127/0941-2948/2007/0256.

Eigbe, U., Beck, M.B., Wheater, H.S., and Hirano, F. 1998. Kalman Filtering in Groundwater Glow Modelling: Problems and Prospects. *Stochastic Hydrology and Hydraulics* **12** (1): 15–32. doi: 10.1007/s004770050007.

Eppstein, M.J. and Dougherty, D.E. 1996. Simultaneous Estimation of Transmissivity Values and Zonation. *Water Resources Research* **32** (11): 3321–3336. doi: 10.1029/96WR02283.

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

Evensen, G. 2003. The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation. *Ocean Dynamics* **53** (4): 343–367. doi: 10.1007/s10236-003-0036-9.

Evensen, G. 2007. *Data Assimilation: The Ensemble Kalman Filter*. Berlin: Springer Verlag.

Evensen, G. and van Leeuwen, P.J. 2000. An Ensemble Kalman Smoother for Nonlinear Dynamics. *Monthly Weather Review* **128** (6): 1852–1867. doi: 10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2.

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

Evensen. G. 2004. Sampling Strategies and Square Root Analysis Schemes for the EnKF. *Ocean Dynamics* **54** (6): 539–560. doi: 10.1007/s10236-004-0099-2.

Floris, F.J.T., Bush, M.D., Cuypers, M., Roggero, F., and Syversveen, A-R. 2001. Methods for Quantifying the Uncertainty of Production Forecasts: A Comparative Study. *Petroleum Geoscience* **7** (Supplement, 1 May): 87–96.

Furrer, R. and Bengtsson, T. 2007. Estimation of High-Dimensional Prior and Posterior Covariance Matrices in Kalman Filter Variants. *Journal of Multivariate Analysis* **98** (2): 227–255. doi: 10.1016/j.jmva.2006.08.003.

Galli, A., Beucher H., Le Loc'h G., Doligez B., and Heresim group. 1993. The Pros and Cons of the Truncated Gaussian Method. In *Geostatistical Simulations*, ed. M. Armstrong and P.A. Dowd, 217-233. Dordrecht, The Netherlands: Quantitative Geology and Geostatistics, Kluwer Academic Publishers.

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

Gaspari, G. and Cohn, S.E. 1999. Construction of Correlation Functions in Two and Three Dimensions (DAO Office Note 96-03R1). *Quart. J. Roy. Meteor. Soc.* **125** (554): 723–757. doi: 10.1002/qj.49712555417.

Gordon, N.J., Salmond, D.J. and Smith, A.F.M. 1993. Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation. *IEE Proceedings-F* **140** (2): 107–113.

Gu, Y. and Oliver, D.S. 2005. History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter. *SPE J.* **10** (2): 217–224. 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. *J. of Energy Resources Technology* **128** (1): 79–87. doi: 10.1115/1.2134735.

Gu, Y. and Oliver, D.S. 2007. An Iterative Ensemble Kalman Filter for Multiphase Fluid Flow Data Assimilation. *SPE J.* **12** (4): 438–446. SPE-108438-PA. doi: 10.2118/108438-PA.

Hacker, J.P., Anderson, J.L., and Pagowski, M. 2007. Improved Vertical Covariance Estimates for Ensemble-Filter Assimilation of Near-Surface Observations. *Monthly Weather Review* **135** (3): 1021–1036. doi: 10.1175/MWR3333.1.

Hamill, T.M., Whitaker, J.S., and Snyder, C. 2001. Distance-Dependent Filtering of Background Error Covariance Estimate in an Ensemble Kalman Filter. *Monthly Weather Review* **129** (11): 2776–2790. doi: 10.1175/1520-0493(2001)129<2776:DDFOBE>2.0.CO;2.

Hantush, M.M. and Marin o, M.A. 1997. Estimation of Spatially Variable Aquifer Hydraulic Properties Using Kalman Filtering. *Journal of Hydraulic Engineering* **123** (11): 1027–1035. doi: 10.1061/(ASCE)0733-9429(1997)123:11(1027).

Hastie, T., Tibshirani, R., and Friedman, J. 2001. *Elements of Statistical Learning*. New York: Series in Statistics, Springer.

Haugen, V., Nævdal, G., Natvik, L.-J., Evensen, G., Berg, A.M., and Flornes, K.M. 2008. History Matching Using the Ensemble Kalman Filter on a North Sea Field Case. *SPE J.* **13** (4): 382–391. SPE-102430-PA. doi: 10.2118/102430-PA.

Horn, R.A. 1990. The Hadamard Product. In *Matrix Theory and Applications*, ed. C.R. Johnson, Vol. 40, 87–170. New York: Proceedings of Symposia in Applied Mathematics, American Mathematical Society.

Hoteit, I. and Pham, D.-T. 2008. A New Approximate Solution of the Optimal Nonlinear Filter for Data Assimilation in Meteorology and Oceanography. *Monthly Weather Review* **136** (1): 317–334. doi: 10.1175/2007MWR1927.1.

Houtekamer, P.L. and Mitchell, H.L. 1998. Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review* **126** (3): 796–811. doi: 10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Houtekamer, P.L. and Mitchell, H.L. 2001. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review* **129** (1): 123–137. doi: 10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Houtekamer, P.L., Mitchell, H.L., Pellerin, G., Buehner, M., Charron, M., Spacek, L., and Hansen, B. 2005. Atmospheric Data Assimilation With an Ensemble Kalman Filter: Results With Real Observations. *Monthly Weather Review* **133** (3): 604–620. doi: 10.1175/MWR-2864.1.

Hu, L.Y., Le Ravalec, M., Blanc, G., Roggero, F., Noetinger, B., Haas, A., and Corre, B. 1999. Reducing Uncertainties in Production Forecasts by Constraining Geological Modeling to Dynamic Data. Paper SPE 56703 presented at the SPE Annual Technical Conference and Exhibition, Houston, 3–6 October. doi: 10.2118/56703-MS.

Jafarpour, B. and McLaughlin, D. 2007a. Efficient Permeability Parameterization With the Discrete Cosine Transform. Paper SPE 106453 presented at the SPE Reservoir Simulation Symposium, Houston, 26–28 February. doi: 10.2118/106453-MS.

Jafarpour, B. and McLaughlin, D.B. 2007b. History Matching With an Ensemble Kalman Filter and Discrete Cosine Parameterization. Paper SPE 108761 presented at the SPE Annual Technical Conference and Exhibition, Anaheim, California, USA, 11–14 November. doi: 10.2118/108761-MS.

Jafarpour, B. and McLaughlin, D. 2009. Estimating Channelized Reservoirs Permeabilities With the Ensemble Kalman Filter: The Importance of Ensemble Design. *SPE J.* **14** (2): 374–388. SPE-108941-PA. doi: 10.2118/108941-PA.

Jazwinski, A.H. 1970. *Stochastic Processes and Filtering Theory*, Vol. 64. San Diego, California: Mathematics in Science and Engineering, Academic Press.

Julier, S., Uhlmann, J., and Durrant-Whyte, H.F. 2000. A New Method for the Nonlinear Transformation of Means and Covariances in Filters and

Estimators. *IEEE Transactions on Automatic Control* **45** (3): 477–482. doi: 10.1109/9.847726.

Kalman, R.E. 1960. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME—Journal of Basic Engineering* **82** (Series D): 35–45.

Kepert, J.D. 2004. On Ensemble Representation of the Observation-Error Covariance in the Ensemble Kalman Filter. *Ocean Dynamics* **54** (6): 561–569. doi: 10.1007/s10236-004-0104-9.

Kitandis, P.K. 1995. Quasi-Linear Geostatistical Theory for Inversing. *Water Resources Research* **31** (10): 2411–2419. doi: 10.1029/95WR01945.

Kivman, G.A. 2003. Sequential Parameter Estimation for Stochastic Systems. *Nonlinear Processes in Geophysics* **10:** 253–259.

Lawson, W.G. and Hansen, J.A. 2004. Implications of Stochastic and Deterministic Filters as Ensemble Based Data Assimilation Methods in Varying Regimes of Error Growth. *Monthly Weather Review* **132** (8): 1966–1981. doi: 10.1175/1520-0493(2004)132<1966:IOSADF>2.0.CO;2.

Leeuwenburgh, O., Evensen, G., and Bertino, L. 2005. The Impact of Ensemble Filter Definition on the Assimilation of Temperature Profiles in the Tropical Pacific. *Quarterly Journal of the Royal Meteorological Society* **131** (613): 3291–3300. doi: 10.1256/qj.05.90.

Lefebvre, T., Bruyninckx, H., and De Schutter, J. 2004. Kalman Filters for Non-Linear Systems: a Comparison of Performance. *International Journal of Control* **77** (7): 639–653. doi: 10.1080/00207170410001704998.

Leng, C.H. and Yeh, H.D. 2003. Aquifer Parameter Identification Using the Extended Kalman filter. *Water Resources Research* **39** (3): 1062. doi: 10.1029/2001WR000840. doi: 10.1029/2001WR000840.

Li, G. and Reynolds, A.C. 2009. Iterative Ensemble Kalman Filters for Data Assimilation. *SPE J.* **14** (3). SPE-109808-PA. doi: 10.2118/109808-PA.

Liang, B., Alpak, F.O., Sepehrnoori, K., and Delshad, M. 2007. A Singular Evolutive Interpolated Kalman Filter for Rapid Uncertainty Quantification. Paper SPE 106170 presented at the SPE Reservoir Simulation Symposium, Houston, 26–28 February. doi: 10.2118/106170-MS.

Lien, M., Berre, I., and Mannseth, T. 2005. Combined Adaptive Multiscale and Level Set Parameter Estimation. *Multiscale Modeling and Simulation* **4** (4): 1349–1372. doi: 10.1137/050623152.

Liu, N. and Oliver, D.S. 2005a. Critical Evaluation of the Ensemble Kalman Filter on History Matching of Geologic Facies. *SPE Res Eval & Eng* **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. *J. Pet. Sci. Eng.* **47** (3–4): 147–161. doi: 10.1016/j.petrol.2005.03.006.

Lødøen, O.P. and Omre, H. 2008. Scale-Corrected Ensemble Kalman Filtering Applied to Production-History Conditioning in Reservoir Evaluation. *SPE J.* **13** (2): 177–194. SPE-111374-PA. doi: 10.2118/111374-PA.

Lorenc, A.C. 2003. The Potential of the Ensemble Kalman Filter for NWP—a Comparison With 4D-Var. *Quarterly Journal of the Royal Meteorological Society* **129** (595): 3183–3203. doi: 10.1256/qj.02.132.

Lorentzen, R.J. and Nævdal, G. 2008. An Iterative Ensemble Kalman Filter. *IEEE Transactions on Automatic Control* (submitted).

Lorentzen, R.J., Berg, A.M., Nævdal, G., and Vefring, E.H. 2006. A New Approach for Dynamic Optimization of Waterflooding Problems. Paper SPE 99690 presented at the Intelligent Energy Conference and Exhibition, Amsterdam, 11–13 April. doi: 10.2118/99690-MS.

Lorentzen, R.J., Fjelde, K.K., Frøyen, J., Lage, A.C.V.M., Nævdal, G., and Vefring, E.H. 2001a. Underbalanced and Low-Head Drilling Operations: Real Time Interpretation of Measured Data and Operational Support. Paper SPE 71384 presented at the SPE Annual Technical Conference and Exhibition, New Orleans, 30 September–3 October. doi: 10.2118/71384-MS.

Lorentzen, R.J., Fjelde, K.K., Frøyen, J., Lage, A.C.V.M., Nævdal, G., and Vefring, E.H. 2001b. Underbalanced Drilling: Real Time Data Interpretation and Decision Support. Paper SPE 67693 presented at the SPE/IADC Drilling Conference, Amsterdam, 27 February–1 March. doi: 10.2118/67693-MS.

Lorentzen, R.J., Nævdal, G., and Lage, A.C.V.M. 2003. Tuning of Parameters in a Two-Phase Flow Model Using an Ensemble Kalman Filter. *International Journal of Multiphase Flow* **29** (8): 1283–1309. doi: 10.1016/S0301-9322(03)00088-0.

Lorentzen, R.J., Nævdal, G., Vallès, B., Berg, A.M., and Grimstad, A.-A. 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.

Maybeck, P.S. 1979. *Stochastic Models, Estimation, and Control Volume 1*. New York: Mathematics in Science and Engineering, Academic Press.

Mezzadri, F. 2007. How to Generate Random Matrices From the Classical Compact Groups. *Notices of the AMS* **54** (5): 592–604.

Moreno, D. and Aanonsen, S.I. 2007. Stochastic Facies Modelling Using the Level Set Method. In *Petroleum Geostatistics 2007, 10–14 September 2007, Cascais, Portugal*, A18. Utrecht, The Netherlands: Extended Abstracts Book, EAGE Publications BV.

Nævdal, G., Brouwer, D.R., and Jansen, J.-D. 2006. Waterflooding Using Closed-Loop Control. *Computational Geosciences* **10** (1): 37–60. doi: 10.1007/s10596-005-9010-6.

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

Nævdal, G., Mannseth, T., and Vefring, E.H. 2002a. Instrumented Wells and Near-Well Reservoir Monitoring Through Ensemble Kalman Filter. *Proc.,* 8th European Conference on the Mathematics of Oil Recovery, Freiberg, Germany, 3–6 September.

Nævdal, G., Mannseth, T., and Vefring, E.H. 2002b. 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.

Nielsen, L. 2006. Reservoir Characterization by a Binary Level Set Method and Adaptive Multiscale Estimation. PhD thesis, University of Bergen, Bergen, Norway.

Oliver, D.S. and Chen, Y. 2009. Improved Initial Sampling for the Ensemble Kalman Filter. Computational Geosciences, 13(1), 13–26. doi: 10.1007/s10596-008-9102-2.

Oliver, D.S., He, N., and Reynolds, A.C. 1996. Conditioning Permeability Fields to Pressure Data. *Proc.,* 5th European Conference on the Mathematics of Oil Recovery (ECMOR V), Leoben, Austria, 3–6 September, 1–11.

Osher, S. and Sethian, J.A. 1988. Fronts Propagating With Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulation. *Journal of Computational Physics* **79** (1): 12–49. doi: 10.1016/0021-9991(88)90002-2.

Ott, E., Hunt, B.R., Szunyogh, I., Zimin, A.V. Kostelich, E.J., Corazza, M., Kalnay, E., Patil, D.J., and Yorke, J.A. 2004. A Local Ensemble Kalman Filter for Atmospheric Data Assimilation. *Tellus A* **56** (5): 415–428. doi: 10.1111/j.1600-0870.2004.00076.x.

Overbeek, K.M., Brouwer, D.R., Naevdal, G., Jansen, J.D., and van Kruijsdijk, C.P.J.W. 2004. Use of a Virtual Asset to Demonstrate Closed-Loop Reservoir Management. *Proc.,* 9th European Conference on the Mathematics of Oil Recovery, Cannes, France, 30 August–2 September.

Pannekoucke, O., Berre, L., and Desroziers, G. 2007. Filtering Properties of Wavelets for Local Background-Error correlations. *Quarterly Journal of the Royal Meteorological Society* **133** (623): 363–379. doi: 10.1002/qj.33.

Pham, D.T. 2001. Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems. *Monthly Weather Review* **129** (5): 1194–1207. doi: 10.1175/1520-0493(2001)129<1194:SMFSDA>2.0.CO;2.

Rao, K.R. and Yip, P. 1990. *Discrete Cosine Transform: Algorithms, Advantages, Applications*. Boston, Massachusetts: Academic Press.

Reynolds, A.C., Zafari, M., and Li, G. 2006. Iterative Forms of the Ensemble Kalman Filter. *Proc.,* 10th European Conference on the Mathematics of Oil Recovery, Amsterdam, 4–7 September, Paper A030.

Richardsen, S.K. and Randen, T. 2005. Mapping 3D Geo-Bodies Based on Level Set and Marching Methods. In *Mathematical Methods and Modelling in Hydrocarbon Exploration and Production*, ed. A. Iske and T. Randen, Vol. 7, Part II, 247–265. Berlin, Germany: Mathematics in Industry, Springer-Verlag.

Rozier, D., Birol, F., Cosme, E., Brasseur, P., Brankart, J.M., and Verron, J. 2007. A Reduced-Order Kalman Filter for Data Assimilation in Physical Oceanography. *SIAM Review* **49** (3): 449–465. doi: 10.1137/050635717.

Sakov, P. and Oke, P.R. 2008. Implications of the Form of the Ensemble Transformation in the Ensemble Square Root Filters. *Monthly Weather Review* **136** (3): 1042–1053. doi: 10.1175/2007MWR2021.1.

Sethian, J.A. 1999. *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, second edition. Cambridge, UK: Cambridge University Press.

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

Skjervheim, J.-A., Ruud, B.O., Aanonsen, S.I., Evensen, G., Aanonsen, S.I., and Johansen, T.A. 2006. Using the Ensemble Kalman Filter with 4D Data To Estimate Properties and Lithology of Reservoir Rocks. *Proc.,* 10th European Conference on the Mathematics of Oil Recovery, Amsterdam, 4–7 September.

Stengel, R.F. 1994. *Optimal Control and Estimation*. Mineola, New York: Dover Publications.

Strebelle, S. 2002. Conditional Simulation of Complex Geological Structures Using Multiple-Point Statistics. *Mathematical Geology* **34** (1): 1–22. doi: 10.1023/A:1014009426274.

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 SPE 109975 presented at the SPE Annual Technical Conference and Exhibition, Anaheim, California, USA, 11–14 November. doi: 10.2118/109975-MS.

Tippett, M.K., Anderson, J.L., Bishop, C.H., Hamill, T.M., and Whitaker, J.S. 2003. Ensemble Square Root Filters. *Monthly Weather Review* **131** (7): 1485–1490. doi: 10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2.

van Leeuwen, P.J. 1999. Comment on "Data Assimilation Using an Ensemble Kalman Filter Technique." *Monthly Weather Review* **127** (6): 1374–1377. doi: 10.1175/1520-0493(1999)127<1374:CODAUA>2.0.CO;2.

Verlaan, M. and Heemink, A.W. 2001. Nonlinearity in Data Assimilation Applications: A Practical Method for Analysis. *Monthly Weather Review* **129** (6): 1578–1589. doi: 10.1175/1520-0493(2001)129<1578:NIDAAA>2.0.CO;2.

Villegas, R., Dorn, O., Moscoso, M., Kindelán, M., and Mustieles, F. 2006a. Simultaneous Characterization of Geological Shapes and Permeability Distributions in Reservoirs Using the Level Set Method. Paper SPE 100291 presented at the SPE Europec/EAGE Annual Conference and Exhibition, Vienna, Austria, 12–15 June. doi: 10.2118/100291-MS.

Villegas, R., Dorn, O., Moscoso, M.A., and Kindelán, M. 2006b. Simultaneous Characterization of Geological Regions and Parameterized Internal Permeability Profiles in History Matching. *Proc.,* 10th European Conference on the Mathematics of Oil Recovery, Amsterdam, 4–7 September, A015.

Wan, E.A. and Nelson, A.T. 2001. The Unscented Kalman Filter. In *Kalman Filtering and Neural Networks*, ed. S. Haykin, Chap. 7, 221–280. New York: Wiley Series on Adaptive and Learning Systems for Signal Processing, Communications, and Control, John Wiley & Sons.

Wang, C., Li, G., and Reynolds, A.C. 2009. Production Optimization in Closed-Loop Reservoir Management. *SPE J.* **14** (3). SPE-109805-PA. doi: 10.2118/109805-PA.

Wang, X., Bishop, C.H., and Julier, S.J. 2004. Which is Better, an Ensemble of Positive-Negative Pairs or a Centered Spherical Simplex Ensemble? *Monthly Weather Review* **132** (7): 1590–1605. doi: 10.1175/1520-0493(2004)132<1590:WIBAEO>2.0.CO;2.

Wen, X.-H. and Chen, W.H. 2006. Real-Time Reservoir Model Updating Using Ensemble Kalman Filter With Confirming Option. *SPE J.* **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. *SPE J.* **12** (2): 156–166. SPE-111571-PA. doi: 10.118/111571-PA.

Whitaker, J.S. and Hamill, T.M. 2002. Ensemble Data Assimilation Without Perturbed Observations. *Monthly Weather Review* **130** (7): 1913–1924. doi: 10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Wu, Z. and Datta-Gupta, A. 2002. Rapid History Matching Using a Generalized Travel-Time Inversion Method. *SPE J.* **7** (2): 113–122. SPE-78359-PA. doi: 10.2118/78359-PA.

Zafari, M. and Reynolds, A.C. 2007. Assessing the Uncertainty in Reservoir Description and Performance Predictions With the Ensemble Kalman Filter. *SPE J.* **12** (3): 382–391. SPE-95750-PA. doi: 10.2118/95750-PA.

Zang, X. and Malanotte-Rizzoli, P. 2003. A Comparison of Assimilation Results From the Ensemble Kalman Filter and a Reduced-Rank Extended Kalman Filter. *Nonlinear Processes in Geophysics* **10** (6): 477–491.

Zhang, D., Lu, Z., and Chen, Y. 2007. Dynamic Reservoir Data Assimilation With an Efficient, Dimension-Reduced Kalman Filter. *SPE J.* **12** (1): 108–117. SPE-95277-PA. doi: 10.2118/95277-PA.

Zhang, F., Reynolds, A.C., and Oliver, D.S. 2003. The Impact of Upscaling Errors on Conditioning a Stochastic Channel to Pressure Data. *SPE J.* **8** (1): 13–21. SPE-83679-PA. doi: 10.2118/83679-PA.

Zhao, Y. 2008. Ensemble Kalman Filter Method for Gaussian and Non-Gaussian Priors. PhD thesis, University of Tulsa, Tulsa, Oklahoma.

Zhao, Y., Reynolds, A.C., and Li, G. 2008. Generating Facies Maps by Assimilating Production Data and Seismic Data With the Ensemble Kalman Filter. Paper SPE 113990 presented at the SPE/DOE Symposium on Improved Oil Recovery, Tulsa, 20–23 April. doi: 10.2118/113990-MS.

Zupanski, M. 2005. Maximum Likelihood Ensemble Filter: Theoretical Aspects. *Monthly Weather Review* **133** (6): 1710–1726. doi: 10.1175/MWR2946.1.

Sigurd Ivar Aanonsen is a principal engineer at StatoilHydro and adjunct professor at the Centre for Integrated Petroleum Research at the University of Bergen. He holds a PhD degree in applied mathematics from the University of Bergen. Aanonsen's research interests include inverse problems and parameter estimation for conditioning reservoir models to dynamic data (production data and 4D seismic), assessment of uncertainty, optimization, and reservoir management. He has 18 years of experience from Norsk Hydro as a reservoir engineer and research scientist and has been a member of the *SPE Journal* editorial board. **Geir Nævdal** is a chief scientist at the International Research Institute of Stavanger and an adjunct principal researcher at the Centre for Integrated Petroleum Research. He holds a PhD degree in mathematics from the University of Trondheim. **Dean S. Oliver** is the Mewbourne Chair Professor of the Mewbourne School of Petroleum and Geological Engineering at the University of Oklahoma. He was director of the department from 2002 to 2006. Before joining the University of Oklahoma, he was a professor in the Petroleum Engineering department at the University of Tulsa for six years and worked 17 years for Chevron as a research geophysicist, a staff reservoir engineer for Chevron USA, for Saudi Aramco, and as a research scientist in reservoir characterization. He has more than 40 publications in refereed journals. He received the SPE Reservoir Description and Dynamics Award in 2004, the Distinguished Member Award in 2008, and is currently the executive editor of *SPE Journal*. **Albert C. Reynolds** is professor of petroleum engineering and mathematics, holder of the McMan Chair in Petroleum Engineering, and Director of TUPREP at The University of Tulsa, where he has been a faculty member since 1970. He has authored or coauthored more than 100 technical papers and two books. His research interests include optimization, assessment of uncertainty, data integration, history matching, and well testing. He holds a BA degree from the University of New Hampshire, an MS degree from Case Institute of Technology, and a PhD degree from Case Western Reserve University, all in mathematics. He received the SPE Reservoir Description and Dynamics Award in 2003, the SPE Formation Evaluation Award in 2005, and became a Distinguished Member in 2001. **Brice Vallès** is a senior research engineer at IRIS in Bergen, Norway, where he has extensively worked on the ensemble Kalman filter since 2004. email: brice.valles@iris.no. Previously, he worked for more than 2 years as a research scientist for Schlumberger Stavanger Research on reservoir model building and 4D history matching. His research interests are in data assimilation (EnKF), reservoir characterization, computational fluid dynamics, and fluid mechanics. Vallès holds a BS degree in mechanics from the University of Aix-Marseille I, an MS degree in mechanics and aeronautics from the University of Aix-Marseille II, a DEA in transfers and fluid mechanics from the Grenoble Institute of Technology, and a PhD in computational fluid dynamics from the Norwegian University of Science and Technology.