This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier’s archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright
Comparison of deterministic ensemble Kalman filters for assimilating hydrogeological data

Alexander Y. Sun*, Alan Morris, Sitakanta Mohanty

Geosciences and Engineering Division, Southwest Research Institute®, 6220 Culebra Road, San Antonio, TX 78238, United States

1. Introduction

Accurate prediction of groundwater availability and fates of subsurface contaminants requires calibrated, high-fidelity groundwater models. Many researchers have attempted the optimization of model structures and parameters in the past [cf., [23,34,39]]. However, two types of inherent difficulties complicate the solution of inverse problems in groundwater modeling: (i) the geological structure of a real aquifer is often complex and unknown and (ii) data that can be used for model calibration are usually limited in both quantity and quality. Groundwater model calibration is traditionally done in a batch-processing manner in which all known data are used at the front to solve an optimization problem. Recently, there has been a renewed interest in sequential data assimilation techniques across a broad range of disciplines including hydrology. The trend is driven not only by new advances in data collection technologies and high-dimensional assimilation schemes, but also by the need to automate decision making for optimal water resources management. It is now widely recognized that reconciling geologic models to the dynamic response of aquifers or oil and gas reservoirs is critical for subsurface characterization and for building reliable models [11].

Data assimilation refers to a process of fusing observations with prior knowledge to obtain updated and, hopefully, improved estimates of the distribution of the true model state or parameters [38]. Observing and analyzing aquifer responses to system excitation, such as pumping events, can reveal important features of the subsurface formation heterogeneity. Although pressure head data have been the de-facto state observations to use in groundwater model calibration, a number of other observable attributes can also be potentially incorporated to infer model parameters via appropriately coupled models [34]. Examples of such data include, for example, tracer concentration and arrival time, water temperature, infiltration rate, water content, and geophysical data. Because the true model is never known exactly and observations are rarely perfect, data assimilation is usually cast in a Bayesian statistical framework including three major components: a forecast model that propagates the system dynamics forward in time, an observation model that maps the observations to model state variables, and probability density functions (PDF) of model and measurement errors.

* Corresponding author.
E-mail address: asun@swri.org (A.Y. Sun).
The dimensions of model states in high-resolution groundwater models can be on the order of $10^5$ to $10^7$. Propagating error statistics in such high-dimensional space is computationally prohibitive if the traditional filtering schemes, such as the Kalman filter or importance-sampling/resampling particle filters, are used without modification. The nonlinearity of system dynamics and the often poorly characterized model error distributions impose additional challenges. An important area of recent research focuses on using reduced-rank assimilation schemes as computationally feasible approximate solutions to the high-dimensional filtering problems.

In his seminal paper, Evensen [16] introduced a Monte-Carlo type, ensemble Kalman filter (EnKF) that approximates system error statistics using an ensemble of model states representing discrete samples of the error distribution. The most attractive feature of the EnKF is that the ensemble size required is usually much smaller than the dimension of model states and, therefore, it can be applied to solving high-dimensional problems which are intractable using the traditional Kalman filter. Another advantage of the EnKF is that it preserves some information about the prior covariance structure between assimilating steps, which can be crucial for stabilizing the filter performance in high-dimensional space [2]. Although in principle the EnKF is optimal only for Gaussian error statistics, the Monte-Carlo nature of the method allows it to provide useful estimates in the case of mildly non-Gaussian PDFs because it does not invoke explicit assumptions (e.g., linearity) about the underlying dynamic model [12, 14, 17]. Of course, the fact that the EnKF only propagates the first two moments of error statistics precludes its effectiveness in the case of highly nonlinear uncertainty evolution (e.g., multimodal distributions). The EnKF has been extensively studied in the fields of oceanography and atmospherics, and numerous variants have been proposed to extend it to increasingly general settings.

Various existing ensemble filters mainly differ in how the ensemble members are updated (i.e., the analysis step) and can thus be categorized as perturbation-based or deterministic filters. Perturbation-based ensemble filters (including the EnKF) add artificial random noise to each observation so that the desired analysis ensemble covariance is replicated [9]. The added observation noise, however, becomes an extra source of inaccuracy. In comparison, deterministic ensemble filters bypass the need for sample perturbation by applying linear algebraic transformations to produce analysis ensembles that match the desired sample mean and covariance exactly. Thus, deterministic ensemble filters tend to be more robust and obtain more accurate estimates than the perturbation-based approaches do, especially for small-sized ensembles [36].

Undoubtedly, the performance of ensemble filters relies on how good the underlying error statistics is represented and propagated. This is of particular concern when ensemble filters are used to solve real-world problems and the assumption of perfect model must be relaxed. Data limitations combined with reduced-rank approximations may lead to incorrect representation of covariance structures and underestimation of model error statistics. Underestimation of the forecast covariance may eventually cause filter divergence, a point at which new observations cease to have impact on the ensemble members and the filter estimates start to drift away from the actual distribution of the model state. Essentially, a core issue to deal with in the analysis step is how to generate new ensemble members to preserve forecast covariance while mitigating the impact of model uncertainty. Covariance inflation and localization are commonly used to counter filter divergence.

Covariance inflation is an ad hoc technique that attempts to address the filter divergence issue by inflating the analysis covariance during each assimilation cycle [2, 8]. Multiplicative inflation simply multiplies the forecast ensemble covariance by a tunable scalar. Additive inflation adds random perturbations with a certain covariance structure to each ensemble member, which is only feasible if the model error can be well characterized. An alternative interpretation for covariance inflation is that it constitutes a regularization technique for inversion under model uncertainty. There is a direct link between covariance inflation and the widely used regularized least squares method for ill-conditioned systems [34]. Covariance inflation can also be tied to robust programming: a hydrogeological application of the latter was presented in [35], where the bound of hydraulic conductivity uncertainty was used to regularize the covariance matrix in a Bayesian-based robust contaminant source identification method.

Localization is another ad hoc technique for stabilizing filter performance [3, 19, 22]. The basic idea behind localization is to restrict the radius of influence of each observation so that a certain observation only affects the state variables that are close to it in the physical space. An alternative explanation is that localization solves for a small model state in a relatively large ensemble space and thus be categorized as perturbation-based or deterministic filters. In hydrogeology, interests in ensemble-based data assimilation techniques are driven by the needs to continuously reduce uncertainties associated with flow and transport models; ensemble-based data assimilation is largely used as a means to learn posterior PDFs of high-dimensional, spatially variable model parameters. The viability of the EnKF has been demonstrated for history matching of petroleum reservoir models (e.g., [5, 18, 25, 37, 40]) and fusion in hydrologic applications (e.g., [24, 29]). Recently, Chen and Zhang [10] demonstrated the capability of EnKF for continuously updating the hydraulic conductivity field in groundwater models through assimilating dynamic data (head) and static data (hydraulic conductivity). They showed that the EnKF is an efficient data assimilation tool and is robust even when the prior estimation of error statistics is slightly biased.

A significant part of recent algorithm development activities in the data assimilation research community focuses on the deterministic ensemble filters. The EnKF has been shown to give good accurate and stable performance. Because all ensemble filters are suboptimal filters, the subtle difference in the analysis scheme may have important implications on filter performance. Even a slight improvement in model prediction capability may have profound economic and environmental impacts. Filters that perform well in one type of problems may not deliver the same performance in others. The superb performance of the deterministic ensemble filters has mostly been demonstrated in the context of oceanography and atmospherics. Therefore, the main purpose of this study is to comparatively study the efficacy of several deterministic ensemble filters for assimilating hydrogeological data. We hope the results can serve as a reference for groundwater modelers to choose the most appropriate deterministic ensemble filter to use. For this purpose, we selected (i) the singular evolution interpolated Kalman filter (SEIK) [21, 28]; (ii) the ensemble transform Kalman filter (ETKF) [8, 36]; (iii) the deterministic ensemble Kalman filter (DEnKF) [32]; and (iv) the local ensemble transform Kalman filter (LEnKF) [22, 27].

SEIK and ETKF are the so-called square root filters [36]; however, they use different matrix decomposition techniques to generate the analysis ensemble. DEnKF was proposed as a simpler alternative to the more complex matrix decomposition schemes.
involved in square root filters. It is similar in form to the EnKF but does not require artificial observation perturbation. Finally, LETKF involves a grid-based localization scheme (i.e., the aforementioned explicit localization technique) so that assimilation can be carried out more efficiently on computer clusters. The main interest in this study is to use LETKF as a vehicle for testing the effectiveness of localization in hydrogeological applications. Note that the deterministic filters mentioned here were mostly exemplified by high-dimensional but densely sampled problems in oceanography or atmospherics.

This paper is organized as follows. Section 2 introduces the background and theoretical formulation of the deterministic filters under consideration. Section 3 compares the performance of the filters using both one- and two-dimensional groundwater flow problems. The one-dimensional problem not only has pedagogic meaning, but also serves as an example of low-dimensional, densely sampled systems. In contrast, the two-dimensional problem is moderately high-dimensional and is sparsely sampled. Finally, discussion and conclusions are given in Section 4.

2. Theoretical formulation of filters

2.1. Background

The governing equation of groundwater flow in saturated aquifers is

\[ \frac{\partial h}{\partial t} = \nabla \cdot (k \nabla h) + J \]  \hspace{1cm} (1)

where \( h \) [L] is pressure head, \( S_s \) [L\(^{-1}\)] is specific storage coefficient, \( k \) is hydraulic conductivity [L/T], and \( J \) [L/T] represents all sink/source terms. Eq. (1) can be solved either numerically or analytically with appropriate initial and boundary conditions.

Now let \( \mathbf{h} \in \mathbb{R}^m \), given in the form of an \( m \)-dimensional column vector, be a discrete representation of the system pressure head at time \( t \); Eq. (1) can be cast into a recursive state estimation problem in which \( \mathbf{h} \) is solved using known model parameters and the solution from the previous step as the initial condition. Because of the spatial heterogeneity, \( k \) is always uncertain and a joint state and parameter estimation problem is often solved. Although the parameter estimation problem can be solved in the context of batch optimization, the sequential parameter estimation approaches become more advantageous when both the quantity and frequency of observations are large, as often encountered in real- or near real-time monitoring. Other advantages of sequential estimation have been discussed in depth in Refs. [14,23]. To accommodate joint state and parameter estimation, a standard practice in system engineering is to augment the state vector with uncertain parameters. Here we simply assume that \( k \) at all grid blocks needs to be updated during each assimilation cycle. Note that (i) it is often better to use a more parsimonious parameterization of the hydraulic conductivity field to reduce the degree of freedom and, thus, alleviate the ill-posedness of the inverse problem and (ii) \( k \) is not a time-varying parameter per se; however, if all other sources of model uncertainties are excluded, the spatial variability in \( k \) becomes the only stochastic forcing factor.

Let \( x_i \in \mathbb{R}^d \) denote the resulting augmented state vector and \( x_{i,t} \) represent the set of all available observations up to \( t \). The prior PDF of the state at \( t \) is

\[ p(x_{i,t}) = \int p(x_{i-1}) | p(x_{i-1} | x_{i-1}) | dx_{i-1} \]  \hspace{1cm} (2)

which can be updated via Bayes' theorem to arrive at the analysis or posterior PDF

\[ p(x_{i,t} | z_{i,t}) = \text{const} \cdot p(z_{i,t}) | p(x_{i}, | z_{i,t}) | \]  \hspace{1cm} (3)

where const is a normalizing constant, \( z_{i,t} \in \mathbb{R}^d \) is the observation vector at \( t \), and \( p(z_{i,t}) \) the likelihood function. The classic Kalman filter is the optimal filter for sequential updating when models are linear and the error distributions are Gaussian

\[ x_i = F x_{i-1} + \omega_i, \omega_i \sim N(0, Q_i) \]  \hspace{1cm} (4)

\[ z_i = H x_i + \epsilon_i, \epsilon_i \sim N(0, R_i) \]  \hspace{1cm} (5)

where \( F \) is a linear model operator that evolves the model state, \( H \) is a linear measurement operator that maps the model prediction to observations, \( N(\cdot, \cdot) \) denotes Gaussian distributions, and \( Q \) and \( R \) are the model error and measurement error covariance, respectively. Under the assumptions of the Kalman filter, minimization of the posterior PDF in Eq. (3) yields the following Kalman analysis scheme

\[ x_i^f = x_i^a + K_i(z_i - H x_i^a) \]  \hspace{1cm} (6)

\[ P_i^f = (I - K_i H_i) P_i^a \]  \hspace{1cm} (7)

where the superscripts \( f \) and \( a \) are used to differentiate the forecast and analysis states, \( P \) represent the state covariance matrices, and \( K \) is the Kalman gain matrix defined as

\[ K_i = P_i^a H_i^T (H_i P_i^a H_i^T + R_i)^{-1} \]  \hspace{1cm} (8)

The Kalman filter is inappropriate for high-dimensional nonlinear systems because of the inherent linearity assumption and high computational cost associated with calculating state covariance matrices. The EnKF partly addresses these limitations by replacing the full state forecast covariance \( P \) required for Eqs. (7) and (8) with a lower-rank, sample covariance matrix calculated from an ensemble of model states. Given an \( N \)-member ensemble, \( x_i = \{ x_i^{1}, \ldots, x_i^{N} \} \), the sample mean and covariance are estimated as

\[ x_i^f = \frac{1}{N} \sum_{i=1}^{N} x_i \]  \hspace{1cm} (9)

\[ P_i^f = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - x_i^f)(x_i - x_i^f)^T = \frac{1}{N-1} A A^T \]  \hspace{1cm} (10)

where \( A \) is an ensemble perturbation matrix with its \( i \)-th column defined by \( A_i = x_i^f - x_i^a \). In practice, one can always calculate the product \( HA \) first so that the more expensive calculation defined in Eq. (10) is never necessary [14]. Use of the approximated ensemble covariance \( P_i^f \) results in an analysis error covariance \( P_i^f \) smaller than the Kalman filter analysis covariance defined in Eq. (7). The EnKF remedy is to perturb the observations using artificial Gaussian sampling noise so that the analysis state becomes

\[ x_i^{a,f} = x_i^{a} + K_i(z_i + d_i - H x_i^{a,f}), \quad i = 1, \ldots, N \]  \hspace{1cm} (11)

where \( d_i \) contains the artificial random noise. As mentioned before, the perturbation-based approach introduces an additional source of inaccuracy in the filter estimates. More specifically, the artificial noise may destroy the information about prior relations between state variables and cause filter divergence. Various deterministic filters introduced in the remainder of the section attempt to circumvent this issue.

2.2. Deterministic ensemble filters

2.2.1. The SEIK filter

SEIK, originally proposed by Pham [28], is a deterministic ensemble filter that uses a preconditioned ensemble and a computationally efficient analysis scheme. In the following, we mainly reference the SEIK implementation described in Refs. [21,26]. SEIK consists of four stages: initialization, forecast, analysis, and reinitialization. A pseudo implementation of SEIK in our demonstration follows:

\[ p(x_{i,t} | z_{i,t}) = \text{const} \cdot p(z_{i,t}) | p(x_{i}, | z_{i,t}) | \]  \hspace{1cm} (3)

where const is a normalizing constant, \( z_{i,t} \in \mathbb{R}^d \) is the observation vector at \( t \), and \( p(z_{i,t}) \) the likelihood function. The classic Kalman filter is the optimal filter for sequential updating when models are linear and the error distributions are Gaussian

\[ x_i = F x_{i-1} + \omega_i, \omega_i \sim N(0, Q_i) \]  \hspace{1cm} (4)

\[ z_i = H x_i + \epsilon_i, \epsilon_i \sim N(0, R_i) \]  \hspace{1cm} (5)

where \( F \) is a linear model operator that evolves the model state, \( H \) is a linear measurement operator that maps the model prediction to observations, \( N(\cdot, \cdot) \) denotes Gaussian distributions, and \( Q \) and \( R \) are the model error and measurement error covariance, respectively. Under the assumptions of the Kalman filter, minimization of the posterior PDF in Eq. (3) yields the following Kalman analysis scheme

\[ x_i^f = x_i^a + K_i(z_i - H x_i^a) \]  \hspace{1cm} (6)

\[ P_i^f = (I - K_i H_i) P_i^a \]  \hspace{1cm} (7)

where the superscripts \( f \) and \( a \) are used to differentiate the forecast and analysis states, \( P \) represent the state covariance matrices, and \( K \) is the Kalman gain matrix defined as

\[ K_i = P_i^a H_i^T (H_i P_i^a H_i^T + R_i)^{-1} \]  \hspace{1cm} (8)

The Kalman filter is inappropriate for high-dimensional nonlinear systems because of the inherent linearity assumption and high computational cost associated with calculating state covariance matrices. The EnKF partly addresses these limitations by replacing the full state forecast covariance \( P \) required for Eqs. (7) and (8) with a lower-rank, sample covariance matrix calculated from an ensemble of model states. Given an \( N \)-member ensemble, \( x_i = \{ x_i^{1}, \ldots, x_i^{N} \} \), the sample mean and covariance are estimated as

\[ x_i^f = \frac{1}{N} \sum_{i=1}^{N} x_i \]  \hspace{1cm} (9)

\[ P_i^f = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - x_i^f)(x_i - x_i^f)^T = \frac{1}{N-1} A A^T \]  \hspace{1cm} (10)

where \( A \) is an ensemble perturbation matrix with its \( i \)-th column defined by \( A_i = x_i^f - x_i^a \). In practice, one can always calculate the product \( HA \) first so that the more expensive calculation defined in Eq. (10) is never necessary [14]. Use of the approximated ensemble covariance \( P_i^f \) results in an analysis error covariance \( P_i^f \) smaller than the Kalman filter analysis covariance defined in Eq. (7). The EnKF remedy is to perturb the observations using artificial Gaussian sampling noise so that the analysis state becomes

\[ x_i^{a,f} = x_i^{a} + K_i(z_i + d_i - H x_i^{a,f}), \quad i = 1, \ldots, N \]  \hspace{1cm} (11)

where \( d_i \) contains the artificial random noise. As mentioned before, the perturbation-based approach introduces an additional source of inaccuracy in the filter estimates. More specifically, the artificial noise may destroy the information about prior relations between state variables and cause filter divergence. Various deterministic filters introduced in the remainder of the section attempt to circumvent this issue.
1. Initialization stage. Generate \( N \) realizations of \( k \) field based on the predetermined statistics. Optionally, the realizations can be conditioned by hard data (e.g., hydraulic conductivity or well logs).

2. Forecast stage.
   - For \( i = 1 : N \), solve for \( \mathbf{h}_i \) and form an augmented state vector \( \mathbf{x}_i^f \) by concatenating \( \mathbf{h}_i \) and all \( k \) values. This results in a \( 2m \times N \) ensemble of augmented state vectors, \( \mathbf{X}_i^f = (\mathbf{x}_i^f)^T \).
   - Factorize the forecast covariance matrix into \( \mathbf{P}_i^f = \mathbf{L}_i \mathbf{G}_i^T \), where
     \[
     \mathbf{L}_i = \mathbf{X}_i^f \quad \text{and} \quad \mathbf{G} = (\mathbf{N}_i^T \mathbf{N}_i)^{-1}
     \]
     in which \( \mathbf{T} \) (cf. Eq. 46 in [26]) is a transformation matrix that implicitly subtracts the sample mean from \( \mathbf{X}_i^f \).

3. Analysis stage.
   - Calculate the Kalman gain matrix
     \[
     \mathbf{K}_i = \mathbf{L}_i \mathbf{U}_i (\mathbf{H}_i \mathbf{L}_i)^T \mathbf{R}_i^{-1}
     \]
     where the inverse of \( \mathbf{U}_i \) is given by
     \[
     \mathbf{U}_i^{-1} = \rho_i \mathbf{N}_i^T + (\mathbf{H}_i \mathbf{L}_i)^T \mathbf{R}_i^{-1} \mathbf{H}_i \mathbf{L}_i
     \]
     in which \( \rho_i < 1 \) is a “forgetting factor” that has similar function as the covariance inflation factor. The analysis ensemble covariance is implicitly given by
     \[
     \mathbf{P}_i^a = \mathbf{L}_i \mathbf{U}_i \mathbf{L}_i^T - \mathbf{K}_i (\mathbf{L}_i - \mathbf{H}_i \mathbf{L}_i^T) \mathbf{K}_i^T
     \]
   - Update the forecast ensemble mean with the latest observations
     \[
     \mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}_i (\mathbf{z}_i - \mathbf{H}_i \mathbf{x}_i^f)
     \]

4. Reinitialization. The \( k \)th member is updated by the following equation to form the analysis ensemble
   \[
   \mathbf{x}_i^a = \mathbf{x}_i^f + \sqrt{\mathbf{K} \mathbf{L}_i} \mathbf{C}_i^{-1} \mathbf{\Omega}_i^a, \quad i = 1, \ldots, N
   \]
   where \( \mathbf{C}_i^{-1} \) comes from the Cholesky decomposition of \( \mathbf{U}_i^{-1} \) and \( \mathbf{\Omega}_i^a \) is a random matrix whose columns are orthonormal and orthogonal to the vector, \( (1, \ldots, 1)^T \). The use of \( \mathbf{\Omega}_i^a \) constitutes one of the SEIK’s most distinct features called the second-order exact sampling. It ensures that sample mean and covariance are preserved whenever new ensembles are generated. An algorithm for generating \( \mathbf{\Omega}_i^a \) is given in Appendix A of [21].

5. Repeat Steps 2–4 for each assimilation cycle.

Note that (i) the most costly part in SEIK analysis is the Cholesky decomposition of \( \mathbf{U}_i^{-1} \) and the inversion of \( \mathbf{C} \), which are both \( O(N^3) \) operations [26]; (ii) the model error \( \omega_i \) is implicitly expressed via the random space function \( k \) in this problem; other sources of model error may be accounted for through covariance inflation; and (iii) the original initialization step described in Ref. [21] calls for performing second-order exact sampling on the initial ensemble.

We will test the effect of initialization in numerical experiments.

2. Analysis.
   - The key idea in ETKF is to apply an ensemble transformation matrix \( \mathbf{T} \) to form the analysis perturbation matrix such that the resulting analysis covariance matches the theoretical Kalman filter covariance. To find \( \mathbf{T} \), first compute eigenvalue decomposition of an \( N \times N \) matrix \( \mathbf{V} \) defined as
     \[
     \mathbf{V} = \left[ I + \frac{1}{N-1} (\mathbf{H}_i \mathbf{A}_i^T \mathbf{R}_i^{-1} (\mathbf{H}_i \mathbf{A}_i)) \right] = \mathbf{A} \mathbf{T} \mathbf{A}^T,
     \]
     where \( I \) is the identity matrix, \( \mathbf{A} \) is a matrix of orthonormal eigenvectors, and \( \mathbf{T} \) contains eigenvalues of \( \mathbf{V} \) in its diagonal. Because eigenvalue decomposition may experience numerical instability, it is worth examining the eigenvalue spectrum at this step. Our tests showed that eliminating insignificant eigenvalues in this step may significantly stabilize the filter and improve final results in some cases. Care must be taken, however, to avoid loss of accuracy when truncating the eigenvalue spectrum.
   - \( \tilde{\mathbf{T}} \) is obtained as the square root of \( \mathbf{V}^{-1} \)
     \[
     \tilde{\mathbf{T}} = \mathbf{A} \sqrt{\mathbf{T}} \mathbf{A}^T
     \]
     Note that Eq. (18) is different from the original transformation matrix given in Ref. [8]. Sakov and Oke [31] referred to the definition of \( \mathbf{T} \) in Eq. (18) as the double-side solution, and they proved that \( \mathbf{T} \) is also a mean-preserving transformation which ensures that no bias is introduced through the analysis step.

2.2.3. The DEnKF
   In contrast to SEIK and ETKF, which are both square root ensemble filters, the DEnKF offers an elegant alternative of the EnKF scheme. It exploits the observation that if the product \( \mathbf{K} \mathbf{H} \) is small, one can asymptotically match the theoretical Kalman filter covariance up to quadratic terms by halving the Kalman gain \( \mathbf{K} \) [32]. After a similar forecast step as in SEIK, the analysis ensemble perturbation of the DEnKF is given by
     \[
     \mathbf{A}_i^a = \mathbf{A}_i^a - \frac{1}{2} \mathbf{K} \mathbf{H} \mathbf{A}_i^f
     \]
     The mean is updated as in Eq. (15). Compared with the square root ensemble filters, the DEnKF is easier to implement and is relatively robust. The approximation in Eq. (19) adds an extra term (i.e., \( 1/4 \mathbf{K}^T \mathbf{K} \)) to the Kalman analysis covariance formula, Eq. (7). Sakov and Oke showed in [32] that DEnKF substantially outperformed the EnKF and converged as well as or even better than the square root filters; they attributed the observed robust performance of DEnKF to the extra term, which is quadratic in \( \mathbf{K} \mathbf{H} \) and positive semidefinite. Although the extra term may be interpreted as an inherent covariance inflation mechanism as suggested in [32], it differs from the ad hoc covariance inflation and localization techniques because its main role is to eliminate the need to artificially perturbing observations as other deterministic ensemble filters do.

2.2.4. The LETKF
   The LETKF was introduced as an efficient assimilation scheme that scales well to high-dimensional problems involving a large number of observations [22,27]. It is based on the hypothesis that the local neighborhood around each grid node in a discretized system behaves like a low-dimensional unstable system, driven only by local dynamics [22]. Therefore, LETKF is a local–local analysis scheme that operates in the model subspace and measurement subspace pertaining to each grid node. The assimilation problem can then be conveniently divided into \( m \) independent subproblems suitable for parallel treatment, where \( m \) is the number of nodes. The most costly part of the LETKF is obtaining the eigen decomposition of the \( N \times N \) local covariance matrix. The presumption is that
the ensemble size $N$ should be relatively small. LETKF has been demonstrated as an efficient scheme for assimilating data into large-scale weathercast models where the ensemble size is usually small [22].

Fig. 1 uses a one-dimensional problem to illustrate the main components involved in an LETKF analysis, where each local analysis solves a low-dimensional ETKF assimilation problem using all nodes located in a local neighborhood. Each local analysis updates the state vector at one assimilation node and the process is repeated for all nodes in a numerical grid. As mentioned previously, we are mainly interested in using LETKF as a platform for testing the effect of localization. A step-by-step implementation of LETKF is not described here because of its close relation to ETKF. Interested readers may refer to Section 3 of Ref. [22] for a detailed description. The most complex part of the LETKF implementation is formulating and maneuvering a local analysis system for each grid node; however, this is also the part that can be parallelized easily.

3. Numerical experiments

3.1. One-dimensional flow problem

Fig. 2 (top plot) shows the configuration of the one-dimensional flow problem. The length of the column is 100 L and is divided uniformly into 1 L intervals. The thickness of the column is 10 L. Constant heads of 100 L and 90 L are imposed at two ends. The specific storage $S$, of the column is $1 \times 10^{-2} L^{-1}$. A pumping well and a recharge well, both operating at a constant rate of $3 \times 10^{-2} [L/T]$, are placed at 30 and 60 L, respectively. The duration of each assimilation step is 0.1 T and the total number of steps is 6. The hydraulic conductivity is a log-normally distributed random space function, and its log-transform $Y$ is characterized by its mean and standard deviation of $-4.0$ and 0.5, respectively. The spatial correlation of $Y$ is of exponential type [30] with correlation length of 10 L. Stochastic realizations of $Y$ are generated through eigenvalue decomposition of the covariance matrix [13]. We conducted the so-called twin experiment to test the performance of the filters. For this purpose, a realization of $Y$ is randomly chosen from the ensemble to represent the “synthetic truth.” Fig. 2 (center plot) shows the distribution of $Y$ along the length of the column. Fig. 2 (bottom plot) shows the temporal evolution of the pressure head in response to pumping and injection. The solution is obtained by formulating and solving a finite-difference system. All tests were performed in MATLAB$^\text{®}$ on a 1 Gb RAM, 2 GHz Intel$^\text{®}$ Duo-Core laptop.

3.1.1. Effect of observation density

Cases 1–3 are used to test the sensitivity of filters considered here to the number of observations in each assimilation cycle. In Case 1 (also referred to as the base case below), head observations are collected every 5 L along the aquifer. The observation error variance is fixed at 0.01 in all cases. The ensemble size, $N$, varies from 50 to 500. The observation density is every 2 L in Case 2 and 10 L in Case 3. The root mean square error (RMSE) is used to measure the difference between the ensemble mean and the synthetic truth.

Figs. 3a–d show the evolution of RMSE for the base case as a function of different ensemble sizes, where the results of EnKF are also plotted for comparison. The RMSE of most filters stabilizes when $N \geq 100$; however, the DEnKF exhibits relatively robust performance even when $N = 50$. Figs. 4a–c show the final assimilated $Y$ for Cases 1–3 for $N = 250$, while Table 1 reports the final RMSE for all cases considered. The smallest RMSE in each case is highlighted. The following general conclusions can be made based on these results: (i) the RMSE of all filters reduces as the quantity of observations available for assimilation increases; (ii) DEnKF gives the most robust performance and seems to work well even with smaller ensembles; and (iii) the results of SEIK and ETKF are similar, as expected, and they generally perform better for larger ensemble sizes and denser observations. Note that the conclusions drawn above are based on a single synthetic truth. We will conduct sensitivity study in Section 3.2.4 to illustrate the effect of different synthetic “truths” on the filter performance and to assert whether our conclusions still hold.

In terms of computational efficiency, DEnKF is the fastest of the three with a running time under one minute for most cases. Both SEIK and ETKF require matrix decomposition, which may become computationally expensive when the ensemble size is greater than the state vector dimension. Although efficient methods (e.g., snapshot method of Ref. [33]) exist to retain only the largest eigenvalues, one must make some compromise between the loss of numerical accuracy caused by dimension reduction and the numerical efficiency. In addition, our experiences show that the sample covariance may be highly ill-conditioned at times and obtaining the different transformation matrices can be challenging for square root filters. The differences in transformation matrix formulations in SEIK and ETKF contribute to the difference in their performance.

3.1.2. Effect of covariance inflation

Model error can lead to underestimation of ensemble covariance which, in turn, may adversely affect filter performance and cause filter divergence in the most severe case. The actual model error statistics, however, are usually poorly known. As a simplified treatment, a covariance forgetting factor ($\rho_u < 1$) is often used in SEIK as a tuning parameter to compensate for the effect of model errors. Equivalently, a covariance inflation factor ($\nu > 1.0$) can be used in the ETKF.

There can be various sources of model errors in hydrogeological problems [34]. To illustrate the effect of covariance inflation, we assume here that the true $Y$ field has a standard deviation of 1.0. We now attempt to recover the true $Y$ field using the base case ensemble from Section 3.1.1, which was generated using a standard deviation of 0.5. Fig. 5 shows the RMSE of SEIK as a function of different values of $\rho_u$. The use of $\rho_u$ generally improves the performance of the filter in this case. If the value of forgetting factor is excessive as in the case of $\rho_u = 0.5$, however, the performance of the filter can actually degrade. A careful examination of the latter case shows that the filter performance is actually improved during the early assimilation times, prompting the need for an adaptive $\rho_u$. Estimation of the adaptive inflation factor has been considered in Ref. [4].

3.1.3. Effect of localization

LETKF is used to demonstrate the effect of localization on the base case. As mentioned before, the local–local analysis scheme behind LETKF involves two types of local neighborhoods, one for model states and the other for observations. The sizes of the two local neighborhoods are set the same in this experiment.
Fig. 6 shows the RMSE of LETKF as a function of different half-widths of the local neighborhood window ($L_w$). ETKF can be considered as a special case of the LETKF in which a global ensemble analysis is carried out. Thus, the solution obtained by ETKF is also plotted in Fig. 6. As $L_w$ increases, the LETKF solutions converge to the ETKF solution at late assimilation steps, as indicated by the decrease of the RMSE. In this problem, the performance of LETKF is only comparable to that of the ETKF when the local neighborhood is large enough, indicating that the local system surrounding each grid node is not entirely driven by local dynamics in this problem. Head observations located outside the local window also seem to have positive impact on parameter estimation. In fact, first-order stochastic perturbation analyses show that the head perturbations tend to correlate over much longer distances than that of $Y$ ([30], p. 99).

The limitations of the current problem are that the flow problem is one-dimensional and the error statistics is assumed Gaussian. Recently, Agbalaka and Oliver [1] applied localization together with EnKF to automatic history matching of facies distributions. They observed that covariance localization mitigated the problem of variance deficiency in the ensembles. Devegowda et al. [12] showed that covariance localization helped maintain the multimodal nature of a permeability distribution, while the EnKF failed to do so. In nonlinear data assimilation problems (e.g., identifying...
multimodal $Y$ distributions), localization is one of a few means to circumvent the filter divergence issue [7]. Thus, while the effectiveness of localization is not demonstrated by the current problem, the result should not discourage the use of localization as effective dimension reduction and stabilization mechanisms for high-dimensional problems.

---

**Fig. 4.** Log conductivity ($Y$) assimilated by SEIK, ETKF, and DEnKF for observation densities of every (a) 2 L (Case 2); (b) 5 L (Case 1); and (c) 10 L (Case 3).

**Table 1**

The RMSE of the filters for different ensemble sizes (column headers) and observation densities (in parenthesis), where the best performer in each column is highlighted.

<table>
<thead>
<tr>
<th></th>
<th>Case 1 (5 L)</th>
<th>Case 2 (2 L)</th>
<th>Case 3 (10 L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>100</td>
<td>250</td>
<td>500</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>250</td>
<td>500</td>
</tr>
<tr>
<td>EnKF</td>
<td>0.1975</td>
<td>0.1795</td>
<td>0.1594</td>
</tr>
<tr>
<td></td>
<td>0.1539</td>
<td>0.1102</td>
<td>0.0944</td>
</tr>
<tr>
<td></td>
<td>0.2791</td>
<td>0.2564</td>
<td>0.2235</td>
</tr>
<tr>
<td>SEIK</td>
<td>0.1845</td>
<td>0.1767</td>
<td>0.1587</td>
</tr>
<tr>
<td></td>
<td>0.1299</td>
<td>0.1025</td>
<td>0.0832</td>
</tr>
<tr>
<td></td>
<td>0.2886</td>
<td>0.2511</td>
<td>0.2250</td>
</tr>
<tr>
<td>ETKF</td>
<td>0.1895</td>
<td>0.1772</td>
<td>0.1591</td>
</tr>
<tr>
<td></td>
<td>0.1311</td>
<td>0.1025</td>
<td>0.0832</td>
</tr>
<tr>
<td></td>
<td>0.2875</td>
<td>0.2508</td>
<td>0.2256</td>
</tr>
<tr>
<td>DEnKF</td>
<td><strong>0.1761</strong></td>
<td><strong>0.1782</strong></td>
<td><strong>0.1692</strong></td>
</tr>
<tr>
<td></td>
<td><strong>0.1293</strong></td>
<td><strong>0.1074</strong></td>
<td><strong>0.1013</strong></td>
</tr>
<tr>
<td></td>
<td><strong>0.2472</strong></td>
<td><strong>0.2323</strong></td>
<td><strong>0.2226</strong></td>
</tr>
</tbody>
</table>

**Fig. 5.** Effect of covariance inflation illustrated using SEIK, where $\rho_u$ is the “forgetting factor” that adjusts the SEIK forecast covariance. The RMSE is shown in the logarithmic scale.
3.2. A two-dimensional flow problem

3.2.1. Problem set-up

The performance of different deterministic filters is further compared via a two-dimensional synthetic example. Fig. 7 shows the problem set-up. The confined aquifer is represented by a $66 \times 66$ uniform grid, with each cell having a size of $2 \times 2 \, \text{L}$. The aquifer has a thickness of $10 \, \text{L}$ and $S_r$ of $1 \times 10^{-5} \, \text{L}^{-1}$. Stochastic realizations of $Y$ are generated using the GSLIB [13] sequential Gaussian simulator, SGSIM, with a mean of $-3.0$ and standard deviation of 0.5. The spatial correlation of $Y$ is characterized by an isotropic exponential covariance function with correlation length of $15 \, \text{L}$. Similar to settings in petroleum reservoir simulations, all boundaries are set to no-flow boundaries and the system dynamics is driven entirely by a five-spot pumping and injection pattern. The initial head distribution is assumed to be $100 \, \text{L}$ everywhere. The total assimilation time is $2.0 \, \text{T}$ and is done in $0.25 \, \text{T}$ intervals. The locations of pumping/injection wells, as well as the observation wells, are marked in Fig. 7. During each assimilation cycle, head observations are taken from the 24 observation wells and five pumping/injection wells. Observations at different assimilation steps are uncorrelated, and the observation error variance is assumed to be fixed at $0.01$.

The flow problem is solved using MODFLOW [20]. Excluding the boundary cells, the dimension of each augmented state vector is $64 \times 64 = 8192$. We carry out two sets of numerical experiments. In the first, an unconditional ensemble of $Y$ is generated,
from which one realization is randomly chosen as the synthetic truth, and the test proceeds with the rest of the realizations. In the second experiment, a conditional ensemble of $Y$ is generated based on samples taken from the true $Y$ field.

3.2.2. Unconditional ensemble and the effect of observations

We tested the filters for two different ensemble sizes, $N = 200$ and 400. Both numbers are much smaller than the state vector dimension. One of the purposes here is to test the efficacy of using deterministic filters with small-sized unconditional ensembles to represent and propagate error PDFs in high-dimensional space. Covariance inflation is not invoked in SEIK and ETKF. While the ensemble size required to achieve satisfactory performance is certainly problem dependent, the actual number is generally expected to be on the order of hundreds to justify the use of ensemble filters in real settings. For example, sensitivity studies conducted by Chen and Zhang [10] using a $40 \times 40$ grid indicate that 200 is a reasonable ensemble size for the EnKF to give satisfactory results.

Fig. 8 shows the final ensemble mean $Y$ fields obtained by the filters for the case of $N = 200$, along with the true $Y$ field and the initial ensemble mean. All four filters capture the dominant features of the true field well. The DEnKF seems to have introduced more smoothness into the final estimated $Y$ field than others, probably caused by the first-order approximation nature of the filter. Figs. 9a and b show the RMSE for the two different ensemble sizes considered here. The RMSE values show the greatest reduction after the first assimilation step and then become relatively flat, which is similar to that demonstrated in Ref. [10]. A side-by-side comparison between Figs. 9a and b indicate that an increase of ensemble size improves filter performance. All four filters show nearly identical performance when $N = 400$. This is encouraging because as the sample size increases, the difference between the filters diminishes. Overall, the SEIK and DEnKF give the best performance in terms of RMSE. In practice, one wants to seek a balance between the ensemble sizes and filter performance. Filters that deliver the best performance at smaller ensemble sizes are always...
preferred because real models are often computationally costly to
run.

The same experiment was repeated using a sparser monitoring
network where only 19 head observations are used. The sparser
monitoring network is shown by the filled markers in Fig. 7. Figs.
10a and b show the results for \( N = 200 \) and 400, respectively. A vi-
sual comparison between Figs. 9 and 10 allows us to make two
observations. First, the results of all filters are affected because of
the reduction of information content per assimilation cycle; sec-
tond, the EnKF appears to be the worst performer. Because only
the monitoring network is varied between Figs. 9 and 10, the latter
observation implies that artificial observation noise added in EnKF
has a greater adverse impact on the filter performance when the
number of observations is relatively small. This is an important
consideration when choosing ensemble filters because the cover-
age of many hydrogeological monitoring networks can be sparse.

In the literature, SEIK has been shown to require smaller ensemble
sizes for the same performance as the EnKF (e.g., [26]). Our re-
sults support this observation. The standard SEIK scheme
recommends that second-order exact sampling be applied in the
initialization stage. We found that initialization of the ensemble
with second-order exact sampling has little effect on the final result
in the unconditional case. This is partly because the \( Y \) fields gener-
ated using SGSIM already honor the underlying statistics very well.

Of the three deterministic filters, the DEnKF is the most robust
one, as shown by its steadily decreasing RMSE curves in all cases.
The ETKF seems to require a larger ensemble size to achieve the
same level of performance as others.

![Fig. 10. Comparison of filter performance for (a) \( N = 200 \), and (b) \( N = 400 \) unconditional ensemble realizations, where a sparse monitoring network (filled markers shown in Fig. 7) is used.](image)

![Fig. 11. Comparison of filter performance for (a) \( N = 200 \) and (b) \( N = 400 \) conditional ensemble realizations.](image)
### 3.2.3. Effect of conditioning

Measurements of hydrogeologic properties (e.g., conductivity) can be incorporated through either directly augmenting the observation vector as it was done in Ref. [10], or creating conditional ensemble sets. The latter approach is pursued here to illustrate the use of hard data on filter performance. We expect conditioning to improve the filter performance because uncertainty in the initial ensemble is reduced compared to the unconditional case. The conditional dataset was generated by sampling the true Y field every other 15 cells along both directions, resulting in a total of 16 conditioning points. SGSIM was then used to generate conditional Y realizations based on these hard data. The numerical test was repeated for \( N = 200 \) and 400 using the same model configuration as that was used for the unconditional case.

Figs. 11a and b show the evolution of RMSE for \( N = 200 \) and 400, respectively. Compared with the unconditional case, the starting RMSE is smaller because more information is fused to constrain the initial ensemble. For \( N = 200 \), subsequent assimilation using the conditional ensemble makes all filters, except for DEnKF, somewhat unstable. An explanation for this phenomenon is that while conditioning introduces more information at the onset of data assimilation, it also reduces the sample diversity and makes the filter performance unstable for small-sized ensembles. As a result, the minimum ensemble size required for an ensemble filter to represent the same conditional PDF increases. This can be seen from Fig. 11b, where the performance of all filters is substantially more stable after \( N \) is increased to 400.

In both cases, only the DEnKF seems to be immune to the instability caused by conditioning to small-sized ensembles. Applying initialization with second-order exact sampling improved the performance of SEIK in this case. Fig. 12 shows the final ensemble means obtained by all filters for \( N = 400 \), as well as the initial ensemble mean.

### 3.2.4. Sensitivity study

So far the performance of filters has been compared through conducting twin experiments using a single, randomly chosen Y field as synthetic truth. A potential question is why this particular

![Fig. 12. Assimilation using N=400 conditional realizations. Shown are the true Y field, the initial ensemble mean before the assimilation, and the final ensemble means obtained by different ensemble filters.](image-url)

![Fig. 13. Comparison of filter performance for (a) N=200 and (b) N=400 based on 16 different synthetic truths. The symbols with lines represent the mean RMSE and the discrete symbols represent the lower and upper limits of RMSE for each filter.](image-url)
synthetic truth, because filter performance tends to vary for different "truths" and ensemble sizes. To investigate the sensitivity of filter performance to different "truths", we repeat the unconditional experiment in Section 3.2.2 using 16 randomly selected Y fields that are statistically similar to the original one (i.e., all generated using the same geostatistical parameters). The unconditional ensemble used is the same as before.

Figs. 13a and b show the mean RMSE of each method (averaged over the 16 cases) and the associated lower and upper variation bounds for \( N = 200 \) and 400, respectively. The number of observations used is 29. The RMSE between the initial ensemble mean and the true field spans a range of about 0.1, reflecting the unconditional nature of the experiment. The RMSE for \( N = 200 \) shows a large variability. We observe that (i) the performance of all filters except the DEnKF was unstable or even degenerated in some cases, resulting in an increase in the mean RMSE for SEIK, ETKF, and EnKF after the second assimilation step; and (ii) the DEnKF achieved the best performance when the ensemble size is relatively small. All filters become more stable when the ensemble size is increased to 400. The mean RMSE of all three deterministic filters are substantially the same in Fig. 13b. Therefore, filter performance is affected by both the ensemble size and the underlying true field. Also, the fixed configuration of monitoring network and pumping/inject event is more effective for identifying some geologic patterns than others because different geologic structures respond differently to the same pumping/injection events. Our sensitivity study reinforces the conclusions from Section 3.1.1 and Section 3.2.2.

4. Summary and conclusions

Ensemble filters are built on the premise that the PDF of system dynamics can be effectively approximated via a small set of ensembles. As a result, direct manipulation of the full-size Kalman forecast covariance matrix is avoided. Sampling from the high-dimensional posterior distribution, however, is nontrivial even for Gaussian distributions. As a result, all ensemble filters are suboptimal filters that can potentially suffer from matrix rank deficiency and sample poverty (i.e., lack of sample variability). In testing the filters, we look for two filter attributes: robustness and accuracy. Although the two are seemingly contradictory goals, we want to identify filters that most appropriately balance the two.

The one- and two-dimensional synthetic experiments are designed to encompass a wide range of scenarios. The results of our numerical experiments indicate that:

- Deterministic ensemble filters are viable tools for assimilating hydrogeological data and generally achieve better performance than the standard EnKF.
- The DEnKF is consistently the most robust filter in all test cases performed and gives the best performance at relatively small ensemble sizes.
- The SEIK may achieve high accuracy. However, it generally requires larger ensemble sizes than the DEnKF does. Initializing the conditional ensembles with second-order exact sampling improves the filter performance.
- Covariance inflation may effectively improve filter performance in the case of SEIK and ETKF.
- Appropriate sizing of the local neighborhood window is crucial to the effectiveness of the LETKF. If not enough head observations are included in the localized ensemble analysis, the effectiveness of the LETKF will be hampered.
- The spatial coverage of observations is as important as, or even more important than, the temporal coverage.
- Conditioning may significantly reduce the RMSE; however, care must be taken to avoid ensemble degeneracy resulting from sample poverty.
- The use of head observations alone may effectively recover large-scale features of aquifers.
- The adverse effect of artificial sampling noise on EnKF becomes more pronounced when observations are sparse. In contrast, deterministic filters do not suffer from such issue.
- Except for the LETKF, the implementation complexities of deterministic ensemble filters are comparable to that of the EnKF. A common numerical difficulty is related to diagonalization of the ensemble covariance matrix where eigenvalue decomposition is often used.

The deterministic ensemble filters presented here work best when the underlying error distributions are Gaussian. A real aquifer may consist of multiple hydrofacies, each having distinctively different statistics, so that the Y distribution is multimodal multigaussian. Our preliminary experiments show that deterministic ensemble filters can be used to solve these problems only when the nonlinearity is mild and the user is interested in recovering the large-scale features of the underlying field. Otherwise, more sophisticated approaches must be used.

Acknowledgments

This project is funded by Southwest Research Institute through its Internal Research and Development Program (R9704). The authors would like to thank all anonymous reviewers for their constructive comments. The first author is also in debt to Drs. I. Hoteli, T. Bengtsson, P. Sakov, and E. Kostelich for providing their insights during the course of this research.

References


